Supplementary Document for
"TSLS and LIML estimators in Panels with Unobserved Shocks"

Giovanni Forchini
UmeåUniversity

Bin Jiang
Monash University

Bin Peng
University of Bath

March 13, 2018

This document contains proofs of all results presented in the paper "TSLS and LIML estimators in Panels with Unobserved Shocks". The proofs are preceded by short review of stable convergence, conditional strong law of large numbers and conditional central limit theorem.

1 Appendix A: Stable convergence, conditional strong law of large numbers and conditional central limit theorem

This section briefly reviews concepts and results which are used in the main part of the paper but are not easily available in the literature. These include the idea of stable convergence and conditional versions of the strong law of large numbers and the central limit theorem.

The notion of stable convergence of a sequence of random variables was introduced by Rényi (1963): the sequence of random variables ξ_i, $i = 1, 2, ...$ defined on a probability space $(\Omega, \mathcal{A}, P)$ is stable if for any event $B \in \mathcal{A}$ with $P(B) > 0$ the conditional distribution of ξ_i given B tends to a limiting distribution, $\lim_{i\to\infty} P(\xi_i < x|B) = F_B(x)$ for every x which is a continuity point of the distribution function $F_B(x)$, written as $\xi_i \to X$ (stably) (Rényi (1963), p. 294). Stable convergence implies convergence in distribution, and it has been shown that many central limit theorems imply stable convergence (e.g., Aldous and Eagleson (1978) and Hall and Heyde (1980)).

In the panel data model considered in the paper, all random variables are defined on the space $(\Omega, \mathcal{A}, P)$ but the conditioning sets we consider are in the sigma-algebra generated by the factors, $\mathcal{F} \subseteq \mathcal{A}$. Thus, we condition on the events in $\mathcal{F} \subseteq \mathcal{A}$. The notion of stable convergence restricted to these sets is denoted using the terminology of Daley and Vere-Jones (1988) as $\mathcal{F}$-stability (see also Kuersteiner and Prucha (2013)).

The basis of conditional laws of large numbers and conditional central limit theorems is the notion of conditional independence (e.g. Chow and Teicher (1997)). Let $\mathcal{C}$ be a σ-algebra of events and $\{\mathcal{C}_n : n \geqslant 1\}$ a sequence of classes of events. The sequence $\{\mathcal{C}_n : n \geqslant 1\}$ is said to be conditionally independent given $\mathcal{F}$ if for all choices of $C_m \in \mathcal{C}_{k_m}$, where $k_i \neq k_j$ for $i \neq j$, $m = 1, 2, ..., n$, and $n = 2, 3, ...$,

$$P(C_1 \cap C_2 \cap .. \cap C_n|\mathcal{F}) = \prod_{i=1}^{n} P(C_i|\mathcal{F}), \quad a.s..$$

A random sequence $\{X_n : n \geqslant 1\}$ is said to be conditionally independent given $\mathcal{F}$ or, briefly, $\mathcal{F}$-independent, if the sequence of classes $\mathcal{C}_n = \sigma(X_n)$, $n \geqslant 1$ is conditionally independent given $\mathcal{F}$. It should be noted that when $\mathcal{F} = \{\emptyset, \Omega\}$ conditional independence reduces to ordinary stochastic independence of random variables.

Conditions under which conditional independence implies unconditional independence are fully discussed by Phillips (1988). Recent discussions are are given by Majerek et al. (2005), Rao (2009) and Roussas (2008).

The following results are conditional versions of the classical strong laws of large numbers and central limit theorems. Their proofs are very close to those of the unconditional theorems and are reported in detail in the supplementary file (they can also be found in the working paper version of this article - see Forchini et al. (2015)). The conditional central limit theorem can also be seen as a special case of Theorem 3 of Eagleson (1975).

Theorem A.1. Let $\{Z_i : 1 \leq i \leq n\}$ be a sequence of $\mathcal{F}$-independent random variables such that $E\left[|Z_i|^{1+\delta}|\mathcal{F}\right] < \Delta$ for some $\delta > 0$, and Δ being $\mathcal{F}$-measurable with $\Delta < \infty$ a.s. Then conditional on $\mathcal{F}$, $\frac{1}{n}\sum_{i=1}^{n}(Z_i - E[Z_i|\mathcal{F}]) \to 0$ a.s.

Theorems A.1 is the conditional version of Corollary 3.9 of White (2001). It is slightly weaker than the conditional strong laws of large numbers by Majerek et al. (2005), Rao (2009).

Theorem A.2. Let $\{Z_i : 1 \leq i \leq n\}$ be a sequence of $\mathcal{F}$-independent random variables with conditional means $E[Z_i|\mathcal{F}]$, conditional variances $\sigma_i^2 = E\left[(Z_i - E[Z_i|\mathcal{F}])^2|\mathcal{F}\right]$, and $E\left[|Z_i|^{2+\delta|\mathcal{F}}\right] < \Delta$ a.s. for $i = 1, 2, \ldots$ and Δ arbitrary $\mathcal{F}$-measurable, where $\Delta < \infty$ a.s. and some $\delta > 0$. If there is η $\mathcal{F}$-measurable such that $\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 > \eta > 0$ a.s., then conditional on $\mathcal{F}$, $\frac{1}{\bar{\sigma}_n\sqrt{n}}\sum_{i=1}^{n}(Z_i - E[Z_i|\mathcal{F}]) \to^D N(0,1)$ a.s.. Moreover, $\frac{1}{\bar{\sigma}_n\sqrt{n}}\sum_{i=1}^{n}(Z_i - E(Z_i|\mathcal{F})) \to^D N(0,1)$ ($\mathcal{F}$-stably).

Theorem A.2 is the conditional version of the Theorem 5.10 of White (2001). It is slightly weaker versions of the results of Prakasa Rao (2009), Grzenda and Zieba (2008), Yuan, Wei and Lei (2014).

2 Appendix B: Proofs of the results in Sections 3 and 4

Lemma B.1. If Assumption 1.i to 1.vi hold, then the following results hold conditional on $\mathcal{F}$ as $N \to \infty$:

1. $\frac{1}{N}\sum_{i=1}^{N} w_i{}'w_i \to W(F_T)$ a.s.;
2. $\frac{1}{N}\sum_{i=1}^{N} x_i{}'x_i \to X(F_T)$ a.s., where
$$X(F_T) = V(F_T) + \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[\Gamma_i{}'F_T{}'F_T\Gamma_i|\mathcal{F}\right] + v(F_T)'F_T\Gamma(F_T) + \Gamma(F_T)'F_T{}'v(F_T);$$
3. $\frac{1}{N}\sum_{i=1}^{N} w_i{}'x_i \to WX(F_T) = w(F_T)F_T\Gamma(F_T) + \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E[w_i{}'v_i|\mathcal{F}]$ a.s.;
4. $\frac{1}{N}\sum_{i=1}^{N} w_i{}'e_i \to w(F_T)'F_T\gamma(F_T)$ a.s.;
5. $\frac{1}{N}\sum_{i=1}^{N} x_i{}'e_i \to \Gamma(F_T)'F_T{}'F_T\gamma(F_T) + v(F_T)'F_T\gamma(F_T)$ a.s.;
6. $\frac{1}{N}\sum_{i=1}^{N} e_i{}'e_i \to \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[\gamma_i{}'F_T{}'F_T\gamma_i|\mathcal{F}\right] + \Sigma_\varepsilon(F_T)$ a.s.;
7. $\frac{1}{N}\sum_{i=1}^{N}(w_i - \bar{w})'(w_i - \bar{w}) \to W^*(F_T)$ a.s. where $W^*(F_T)$ is defined in (18);
8. $\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})'(x_i - \bar{x}) \to X^*(F_T)$ a.s., where $X^*(F_T)$ is defined in (19);
9. $\frac{1}{N}\sum_{i=1}^{N}(w_i - \bar{w})'(x_i - \bar{x}) \to WX^*(F_T)$ a.s. ,where $WX^*(F_T)$ is defined in (20);
10. $\frac{1}{N}\sum_{i=1}^{N}(w_i - \bar{w})'(e_i - \bar{e}) \to 0$ a.s.;
11. $\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})'(e_i - \bar{e}) \to 0$ a.s.;
12. $\frac{1}{N}\sum_{i=1}^{N}(e_i - \bar{e})'(e_i - \bar{e}) \to \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[\gamma_i{}'F_T{}'F_T\gamma_i|\mathcal{F}\right] - \gamma(F_T)'F_T{}'F_T\gamma(F_T) + \Sigma_\varepsilon(F_T)$ a.s.

The proof of Lemma B.1 is fairly standard so for the sake of simplicity is not reported here. It is available in the supplementary file and in the working paper version of this article.

Proof of Lemma 1.

To prove the first part of the lemma, notice that

$$\hat{\Pi} = \left(\sum_{i=1}^{N}(z_i - \bar{z})'(z_i - \bar{z})\right)^{-1}\left(\sum_{i=1}^{N}(z_i - \bar{z})'(y_i - \bar{y})\right)$$

$$= \Pi + \left(S' \frac{1}{N} \sum_{i=1}^{N} \left(\begin{array}{cc} (w_i - \bar{w})' (w_i - \bar{w}) & (w_i - \bar{w})' (x_i - \bar{x}) \\ (x_i - \bar{x})' (w_i - \bar{w}) & (x_i - \bar{x})' (x_i - \bar{x}) \end{array} \right) S \right)^{-1} S' \left(\frac{1}{N} \sum_{i=1}^{N} \left(\begin{array}{c} (w_i - \bar{w})' (e_i - \bar{e}) \\ (x_i - \bar{x})' (e_i - \bar{e}) \end{array} \right) \right).$$

Thus $\hat{\Pi} \to \Pi$ a.s. from 7-11 of Lemma A.1, where

$$\frac{1}{N} \sum_{i=1}^{N} \left(\begin{array}{cc} (w_i - \bar{w})' (w_i - \bar{w}) & (w_i - \bar{w})' (x_i - \bar{x}) \\ (x_i - \bar{x})' (w_i - \bar{w}) & (x_i - \bar{x})' (x_i - \bar{x}) \end{array} \right) \to Q^* (F_T) = \left(\begin{array}{cc} W^* (F_T) & WX^* (F_T) \\ WX^* (F_T)' & X^* (F_T) \end{array} \right) a.s.$$

conditional on $\mathcal{F}$.

To prove the second part let

$$\begin{aligned}
& \sqrt{N} vec \left(\hat{\Pi} - \Pi \right) \\
= \;& vec \left(\left(S' \frac{1}{N} \sum_{i=1}^{N} \left(\begin{array}{cc} (w_i - \bar{w})' (w_i - \bar{w}) & (w_i - \bar{w})' (x_i - \bar{x}) \\ (x_i - \bar{x})' (w_i - \bar{w}) & (x_i - \bar{x})' (x_i - \bar{x}) \end{array} \right) S \right)^{-1} S' \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\begin{array}{c} (w_i - \bar{w})' (e_i - \bar{e}) \\ (x_i - \bar{x})' (e_i - \bar{e}) \end{array} \right) \right) \\
= \;& \left(I_{p+1} \otimes \left(S' \frac{1}{N} \sum_{i=1}^{N} \left(\begin{array}{cc} (w_i - \bar{w})' (w_i - \bar{w}) & (w_i - \bar{w})' (x_i - \bar{x}) \\ (x_i - \bar{x})' (w_i - \bar{w}) & (x_i - \bar{x})' (x_i - \bar{x}) \end{array} \right) S \right)^{-1} S' \right) \times \\
& \quad \frac{1}{\sqrt{N}} \sum_{i=1}^{N} vec \left(\begin{array}{c} (w_i - \bar{w})' (e_i - \bar{e}) \\ (x_i - \bar{x})' (e_i - \bar{e}) \end{array} \right).
\end{aligned}$$

The term

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} vec \left(\begin{array}{c} (w_i - \bar{w})' (e_i - \bar{e}) \\ (x_i - \bar{x})' (e_i - \bar{e}) \end{array} \right)$$

can be written as

$$\begin{aligned}
& \frac{1}{\sqrt{N}} \sum_{i=1}^{N} vec \left[\left(\begin{array}{c} w_i' - \frac{1}{N} \sum\limits_{j=1}^{N} E[w_j | \mathcal{F}]' \\ x_i' - \frac{1}{N} \sum\limits_{j=1}^{N} E[v_j | \mathcal{F}]' - \Gamma(F_T)' F_T' \end{array} \right) (e_i - F_T \gamma (F_T)) \right] \\
& - vec \left[\left(\begin{array}{c} \bar{w}' - \frac{1}{N} \sum\limits_{j=1}^{N} E[w_j | \mathcal{F}]' \\ \bar{x}' - \frac{1}{N} \sum\limits_{j=1}^{N} E[v_j | \mathcal{F}]' - \Gamma(F_T)' F_T' \end{array} \right) \sqrt{N} (\bar{e} - F_T \gamma (F_T)) \right].
\end{aligned}$$

We will now show that the last term can be neglected. Conditional on $\mathcal{F}$, $\bar{w} - \frac{1}{N} \sum_{i=1}^{N} E[w_i | \mathcal{F}] \to 0$ a.s. and $\bar{x} - \frac{1}{N} \sum_{i=1}^{N} E[v_i | \mathcal{F}] - F_T \Gamma(F_T) \to 0$ a.s. So we just need to prove that $\sqrt{N} \left(\bar{e} - F_T' \gamma (F_T) \right)$ converges to a random matrix conditional on $\mathcal{F}$. Let $\kappa_i = vec(e_i - F_T \gamma(F_T))$ and ζ be an arbitrary $T(p+1) \times 1$ vector. We focus on $\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \zeta' \kappa_i$ and will now verify that the Liapounov's conditions given in Theorem A.2 holds a.s. conditional on $\mathcal{F}$. Notice that

$$E[\zeta' \kappa_i | \mathcal{F}] = 0 \text{ a.s. and } E\left[(\zeta' \kappa_i)^2 | \mathcal{F} \right] = \zeta' E\left[vec(e_i - F_T \gamma(F_T)) \, vec(e_i - F_T \gamma(F_T))' | \mathcal{F} \right] \zeta \quad a.s.$$

Moreover,

$$\begin{aligned}
E\left[|\zeta' \kappa_i|^{2+\delta} | \mathcal{F} \right] &\leqslant |\zeta|_2^{2+\delta} E\left[|vec(e_i - F_T \gamma(F_T))|_2^{2+\delta} | \mathcal{F} \right] = |\zeta|_2^{2+\delta} E\left[|e_i - F_T \gamma(F_T)|_2^{2+\delta} | \mathcal{F} \right] \\
&\leqslant |\zeta|_2^{2+\delta} E\left[\left(|\varepsilon_i|_2 + |F_T (\gamma_i - \gamma(F_T))|_2 \right)^{2+\delta} | \mathcal{F} \right] \\
&\leqslant |\zeta|_2^{2+\delta} 2^{1+\delta} \left(E\left[|\varepsilon_i|_2^{2+\delta} | \mathcal{F} \right] + |F_T|_2^{2+\delta} E\left[|\gamma_i - \gamma(F_T)|_2^{2+\delta} | \mathcal{F} \right] \right),
\end{aligned}$$

where the terms in the last expectations are bounded uniformly a.s. Notice that $\frac{1}{N}\sum_{i=1}^{N} E\left[\left|\zeta'\kappa_i\right|^2|\mathcal{F}\right]$ is a.s. convergent because each term can be uniformly bounded from above a.s. Therefore, using the Cramer-Wold device and Theorem A.2, conditional on $\mathcal{F}$, $\sqrt{N}vec\left(\bar{e}-F_T{}'\gamma\left(F_T\right)\right)$ converges to a random vector and

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N} vec\begin{pmatrix} (w_i-\bar{w})'(e_i-\bar{e}) \\ (x_i-\bar{x})'(e_i-\bar{e})\end{pmatrix} \\
= \; & \frac{1}{\sqrt{N}}\sum_{i=1}^{N} vec\left[\begin{pmatrix} w_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[w_j|\mathcal{F}\right]' \\ x_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[v_j|\mathcal{F}\right]' - \Gamma\left(F_T\right)' F_T{}'\end{pmatrix}\left(e_i - F_T\gamma\left(F_T\right)\right)\right] + o_p(1). \\
= \; & \frac{1}{\sqrt{N}}\sum_{i=1}^{N} vec\left[\begin{pmatrix} w_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[w_j|\mathcal{F}\right]' \\ x_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[v_j|\mathcal{F}\right]' - \Gamma\left(F_T\right)' F_T{}'\end{pmatrix}\left(e_i - F_T\gamma\left(F_T\right)\right)\right] + o_p(1).
\end{aligned}
$$

Let ζ be an arbitrary $(k_1+k_2)(p+1)\times 1$ vector and

$$
\varpi_i = \zeta' vec\left[\begin{pmatrix} w_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[w_j|\mathcal{F}\right]' \\ x_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[v_j|\mathcal{F}\right]' - \Gamma\left(F_T\right)' F_T{}'\end{pmatrix}\left(e_i - F_T\gamma\left(F_T\right)\right)\right].
$$

We can write

$$
\frac{1}{\sqrt{N}}\sum_{i=1}^{N} \zeta' vec\begin{pmatrix} (w_i-\bar{w})'(e_i-\bar{e}) \\ (x_i-\bar{x})'(e_i-\bar{e})\end{pmatrix} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\varpi_i + o_p(1).
$$

Notice that $E\left[\varpi_i|\mathcal{F}\right]=0$ and that $E\left[\left|\varpi_i\right|^{2+\delta}|\mathcal{F}\right] < \Delta$ a.s. is bounded uniformly a.s., since

$$
\begin{aligned}
E\left[\left|\varpi_i\right|^{2+\delta}|\mathcal{F}\right] &\leqslant \left|\zeta\right|_2{}^{2+\delta} E\left[\left|\begin{pmatrix} w_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[w_j|\mathcal{F}\right]' \\ x_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[v_j{}'|\mathcal{F}\right] - \Gamma\left(F_T\right)' F_T{}'\end{pmatrix}\left(e_i - F_T\gamma\left(F_T\right)\right)\right|_2^{2+\delta} |\mathcal{F}\right] \\
\leq \; & \left|\zeta\right|_2{}^{2+\delta} E\left[\left(\left|w_i - \frac{1}{N}\sum_{j=1}^{N} E\left[w_j|\mathcal{F}\right]\right|_2 + \left|x_i - \frac{1}{N}\sum_{j=1}^{N} E\left[v_j|\mathcal{F}\right] - F_T\Gamma\left(F_T\right)\right|_2\right)^{2+\delta} \left|e_i - F_T\gamma\left(F_T\right)\right|_2{}^{2+\delta}|\mathcal{F}\right] \\
\leq \; & \left|\zeta\right|_2{}^{2+\delta} O(1)\, E\left[\left|e_i - F_T\gamma\left(F_T\right)\right|_2{}^{2+\delta}|\mathcal{F}\right] \\
& \cdot\left(E\left[\left|w_i\right|_2{}^{2+\delta}|\mathcal{F}\right] + \left|\frac{1}{N}\sum_{i=1}^{N} E\left[w_i|\mathcal{F}\right]\right|_2^{2+\delta} + E\left[\left|x_i\right|_2{}^{2+\delta}|\mathcal{F}\right] + \left|F_T\Gamma\left(F_T\right)\right|_2{}^{2+\delta} + \left|\frac{1}{N}\sum_{i=1}^{N} E\left[v_i|\mathcal{F}\right]\right|_2^{2+\delta}\right).
\end{aligned}
$$

Since each term above is bounded uniformly, we can conclude that conditional on $\mathcal{F}$

$$
\left(\frac{1}{N}\sum_{i=1}^{N} E\left[\varpi_i{}^2|\mathcal{F}\right]\right)^{-1/2} \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\varpi_i \rightarrow^{D} N(0,1)
$$

so that by using Cramer-Wold device, conditional on $\mathcal{F}$,

$$
\frac{1}{\sqrt{N}}\sum_{i=1}^{N} vec\left[\begin{pmatrix} w_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[w_j|\mathcal{F}\right]' \\ x_i{}' - \frac{1}{N}\sum_{j=1}^{N} E\left[v_j|\mathcal{F}\right]' - \Gamma\left(F_T\right)' F_T{}'\end{pmatrix}\left(e_i - F_T\gamma\left(F_T\right)\right)\right] \rightarrow^{D} \left(\Theta\left(F_T\right)\right)^{\frac{1}{2}} N\left(0, I_{(h_1+h_2+h_3+h_4)(p+1)}\right),
$$

where $\Theta\left(F_T\right)$ is defined as (21).

Proof of Lemma 2.

Write

$$\hat{\Pi} = \Pi + \left(S' \frac{1}{N} \sum_{i=1}^{N} \begin{pmatrix} (w_i - \bar{w})' (w_i - \bar{w}) & (w_i - \bar{w})' (x_i - \bar{x}) \\ (x_i - \bar{x})' (w_i - \bar{w}) & (x_i - \bar{x})' (x_i - \bar{x}) \end{pmatrix} S \right)^{-1} S' \left(\frac{1}{N} \sum_{i=1}^{N} \begin{pmatrix} (w_i - \bar{w})' (e_i - \bar{e}) \\ (x_i - \bar{x})' (e_i - \bar{e}) \end{pmatrix} \right).$$

We have shown that, conditional on $\mathcal{F}$, the matrix in the inverse converges a.s. to $S'Q^* (F_T) S$ and that $\frac{1}{N} \sum_{i=1}^{N} (w_i - \bar{w})' (e_i - \bar{e}) \to 0$ a.s. The remaining term is $\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})' (e_i - \bar{e}) = \frac{1}{N} \sum_{i=1}^{N} x_i{}' e_i - \bar{x}'\bar{e}$. The result follows from the fact that $\bar{e} \to F_T \gamma (F_T)$ a.s., $\bar{x} - F_T \Gamma (F_T) - \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E [v_i | \mathcal{F}] \to 0$ a.s. and $\frac{1}{N} \sum_{i=1}^{N} x_i{}' e_i - \frac{1}{N} \sum_{i=1}^{N} E [v_i{}' | \mathcal{F}] F_T \gamma (F_T) - \frac{1}{N} \sum_{i=1}^{N} E \left[\Gamma_i{}' F_T{}' F_T \gamma_i | \mathcal{F} \right] \to 0$ a.s.

Proof of Theorem 1.

For the panel TSLS estimator of the structural parameters, we have

$$\hat{\beta}_{TSLS} - \beta_0 = \left(\hat{\Pi}_{22}' \hat{H} \hat{\Pi}_{22} \right)^{-1} \hat{\Pi}_{22}' \hat{H} \left(\hat{\pi}_{12} - \hat{\Pi}_{22} \beta_0 \right).$$

It follows from Lemma 1 that conditional on $\mathcal{F}$, $\hat{\Pi}_{22} \to \Pi_{22}$ a.s. and $\hat{\pi}_{12} \to \pi_{12}$ a.s. Moreover, in the proof of Lemma 1, we have shown that

$$\begin{aligned}
\frac{1}{N} \sum_{i=1}^{N} (z_{2,i} - \bar{z}_2)' (z_{2,i} - \bar{z}_2) &\to \underline{S}_2{}' Q^* (F_T) \underline{S}_2 \quad a.s., \\
\frac{1}{N} \sum_{i=1}^{N} (z_{1,i} - \bar{z}_1)' (z_{1,i} - \bar{z}_1) &\to \underline{S}_1{}' Q^* (F_T) \underline{S}_1 \quad a.s., \\
\frac{1}{N} \sum_{i=1}^{N} (z_{2,i} - \bar{z}_2)' (z_{1,i} - \bar{z}_1) &\to \underline{S}_2{}' Q^* (F_T) \underline{S}_1 \quad a.s.
\end{aligned}$$

Thus, $\hat{H} \to H (F_T)$ a.s. follows immediately.

Notice that the convergence is uniform since it does not involve any of the parameters of the model. The first result follows noticing that under Assumption 2, $\pi_{12} = \Pi_{22} \beta_0$.

Now, consider the following term conditional on $\mathcal{F}$,

$$\begin{aligned}
& \sqrt{N} \left(\hat{\pi}_{12} - \hat{\Pi}_{22} \beta_0 \right) = \sqrt{N} \left(\hat{\pi}_{12} - \pi_{12}, \hat{\Pi}_{22} - \Pi_{22} \right) \begin{pmatrix} 1 \\ -\beta_0 \end{pmatrix} \\
= \quad & \sqrt{N} (0, I_{k_2}) \left(\hat{\Pi} - \Pi \right) \begin{pmatrix} 1 \\ -\beta_0 \end{pmatrix} \\
= \quad & \left((1, -\beta_0{}') \otimes (0, I_{k_2}) \right) \sqrt{N} \mathrm{vec} \left(\hat{\Pi} - \Pi \right) \\
\to^D \quad & \left((1, -\beta_0{}') \otimes (0, I_{k_2}) \right) \left(I_{p+1} \otimes \left(S'Q^* (F_T) S \right)^{-1} S' \right) \left(\Theta (F_T) \right)^{\frac{1}{2}} N \left(0, I_{(h_1+h_2+h_3+h_4)(p+1)} \right) \\
= \quad & \left((1, -\beta_0{}') \otimes \left((0, I_{k_2}) \left(S'Q^* (F_T) S \right)^{-1} \right) S' \right) \left(\Theta (F_T) \right)^{\frac{1}{2}} N \left(0, I_{(h_1+h_2+h_3+h_4)(p+1)} \right).
\end{aligned}$$

Thus, conditional on $\mathcal{F}$, $\sqrt{N} \left(\hat{\beta}_{TSLS} - \beta_0 \right) \to^D A (F_T) N \left(0, I_{(h_1+h_2+h_3+h_4)(p+1)} \right)$ follows immediately. This completes the proof of the first part of the theorem.

The proof of the second part of the theorem is established in three steps using the Argmax Theorem (e.g. Theorem 3.2.2 of van der Vaart and Wellner (1996, p. 286). and Theorem 1 of Stock and Wright (2000)). First we establish the consistency of the LIML estimator for β_0. Then, we establish the rate of convergence for the LIML estimator, and finally we show that a rescaled version of the criterion function converges in distribution to a limit process in the space of all uniformly bounded real functions on a compact set for any compact set B.

We have shown that $\hat{\Pi} \to \Pi$ a.s. and $\hat{H} \to H$ a.s. uniformly, so that

$$\left(\hat{\pi}_{12}, \hat{\Pi}_{22} \right)' \hat{H} \left(\hat{\pi}_{12}, \hat{\Pi}_{22} \right) \to (\pi_{12}, \Pi_{22})' H (\pi_{12}, \Pi_{22}) \quad a.s.$$

uniformly. Moreover, $\hat{\Omega} \to \Omega(F_T)$ a.s. by noting results 7-12 of Lemma B.1. Notice also that the convergence is uniform in all the parameters. Thus,

$$
\begin{aligned}
L_N(\beta) &= \frac{(1,-\beta')\left(\hat{\pi}_{12},\hat{\Pi}_{22}\right)'\hat{H}\left(\hat{\pi}_{12},\hat{\Pi}_{22}\right)(1,-\beta')'}{(1,-\beta')\hat{\Omega}(1,-\beta')'} \\
&\to \frac{(1,-\beta')(\pi_{12},\Pi_{22})'H(\pi_{12},\Pi_{22})(1,-\beta')'}{(1,-\beta')\Omega(F_T)(1,-\beta')'} = L_0(\beta) \quad a.s.
\end{aligned}
$$

uniformly. Notice that $L_0(\beta)$ is uniquely minimized at β_0 and it is continuous. Thus, for β in a compact set, Theorem 2.1 of Newey and McFadden (1994) implies that $\hat{\beta}_{LIML} \to \beta_0$ a.s.

Note that the LIML estimator minimizes

$$
\begin{aligned}
L_N(b) &= \frac{(1,-b')\left(\hat{\pi}_{12},\hat{\Pi}_{22}\right)'\hat{H}^{1/2}M_{\hat{H}^{1/2}\hat{\Pi}_{22}}\hat{H}^{1/2}\left(\hat{\pi}_{12},\hat{\Pi}_{22}\right)(1,-b')'}{(1,-b')\hat{\Omega}(1,-b')'} \\
&+\frac{(1,-b')\left(\hat{\pi}_{12},\hat{\Pi}_{22}\right)'\hat{H}^{1/2}P_{\hat{H}^{1/2}\hat{\Pi}_{22}}\hat{H}^{1/2}\left(\hat{\pi}_{12},\hat{\Pi}_{22}\right)(1,-b')'}{(1,-b')\hat{\Omega}(1,-b')'}
\end{aligned}
$$

where $P_Z = Z(Z'Z)^{-1}Z'$ and $M_Z = I_m - P_Z$ are idempotent and projection matrices generated by any $m \times l$ full column rank matrix Z. Since

$$
\begin{aligned}
& (1,-b')\left(\hat{\pi}_{12},\hat{\Pi}_{22}\right)'\hat{H}^{1/2}P_{\hat{H}^{1/2}\hat{\Pi}_{22}}\hat{H}^{1/2}\left(\hat{\pi}_{12},\hat{\Pi}_{22}\right)(1,-b')' \\
= {} & (1,-b')\begin{pmatrix} \hat{\beta}'_{TSLS} \\ I_p \end{pmatrix}\hat{\Pi}'_{22}\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{TSLS}, I_p\right)(1,-b')' \\
= {} & \left(\hat{\beta}_{TSLS}-b\right)'\hat{\Pi}'_{22}\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{TSLS}-b\right),
\end{aligned}
$$

we can write

$$
L_N(b) = \frac{\hat{\pi}'_{12}\hat{H}^{1/2}M_{\hat{H}^{1/2}\hat{\Pi}_{22}}\hat{H}^{1/2}\hat{\pi}_{12} + \left(\hat{\beta}_{TSLS}-b\right)'\hat{\Pi}'_{22}\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{TSLS}-b\right)}{(1,-b')\hat{\Omega}(1,-b')'} \tag{B.1}
$$

Since the LIML estimator minimizes (B.1), we must have

$$
\begin{aligned}
0 &\geqslant N\left(L_N\left(\hat{\beta}_{LIML}\right) - L_N(\beta_0)\right)\left(1,-\hat{\beta}'_{LIML}\right)\hat{\Omega}\left(1,-\hat{\beta}'_{LIML}\right)' \\
&= N\left(\hat{\beta}_{LIML}-\beta_0\right)'\hat{\Pi}'_{22}\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{LIML}-\beta_0\right) \\
&\quad -2N\left(\hat{\beta}_{TSLS}-\beta_0\right)'\hat{\Pi}'_{22}\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{LIML}-\beta_0\right) \\
&\quad +N\left(\hat{\beta}_{TSLS}-\beta_0\right)'\hat{\Pi}'_{22}\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{TSLS}-\beta_0\right)\left(1-\frac{\left(1,-\hat{\beta}'_{LIML}\right)\hat{\Omega}\left(1,-\hat{\beta}'_{LIML}\right)'}{\left(1,-\beta_0{}'\right)\hat{\Omega}\left(1,-\beta_0{}'\right)'}\right) \\
&\quad +N\hat{\pi}'_{12}\hat{H}^{1/2}M_{\hat{H}^{1/2}\hat{\Pi}_{22}}\hat{H}^{1/2}\hat{\pi}_{12}\left(1-\frac{\left(1,-\hat{\beta}'_{LIML}\right)\hat{\Omega}\left(1,-\hat{\beta}'_{LIML}\right)'}{\left(1,-\beta_0{}'\right)\hat{\Omega}\left(1,-\beta_0{}'\right)'}\right).
\end{aligned}
$$

Since $\hat{\beta}_{LIML} \to \beta_0$ a.s., it also converges in probability and we have

$$
1-\frac{\left(1,-\hat{\beta}'_{LIML}\right)\hat{\Omega}\left(1,-\hat{\beta}'_{LIML}\right)'}{\left(1,-\beta_0{}'\right)\hat{\Omega}\left(1,-\beta_0{}'\right)'} = o_p(1),
$$

$$N\left(\hat{\beta}_{TSLS}-\beta_0\right)'\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{TSLS}-\beta_0\right)=O_p(1),$$
$$N\hat{\pi}_{12}'\hat{H}^{1/2}M_{\hat{H}^{1/2}\hat{\Pi}_{22}}\hat{H}^{1/2}\hat{\pi}_{12}=O_p(1).$$

Thus,

$$N\left(\hat{\beta}_{LIML}-\beta_0\right)'\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{LIML}-\beta_0\right)-2N\left(\hat{\beta}_{TSLS}-\beta_0\right)'\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{LIML}-\beta_0\right)+o_p(1)\leqslant 0.$$

Let $\lambda_m\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)$ be the smallest eigenvalue of $\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}$, and notice that

$$N\left(\hat{\beta}_{LIML}-\beta_0\right)'\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{LIML}-\beta_0\right)\geqslant\lambda_m\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)\left\|\sqrt{N}\left(\hat{\beta}_{LIML}-\beta_0\right)\right\|_2^2$$

and

$$\begin{aligned}&\left|N\left(\hat{\beta}_{LIML}-\beta_0\right)'\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{TSLS}-\beta_0\right)\right|\\ \leqslant\;&\left\|\sqrt{N}\left(\hat{\beta}_{LIML}-\beta_0\right)\right\|_2\left\|\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right\|_2\left\|\sqrt{N}\left(\hat{\beta}_{TSLS}-\beta_0\right)\right\|_2.\end{aligned}$$

Using these two inequalities we obtain

$$\begin{aligned}0\;\geqslant\;&N\left(\hat{\beta}_{LIML}-\beta_0\right)'\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{LIML}-\beta_0\right)\\ &-2N\left(\hat{\beta}_{TSLS}-\beta_0\right)'\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\left(\hat{\beta}_{LIML}-\beta_0\right)+o_p(1)\\ \geqslant\;&\lambda_m\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)\left\|\sqrt{N}\left(\hat{\beta}_{LIML}-\beta_0\right)\right\|_2^2\\ &-2\left\|\sqrt{N}\left(\hat{\beta}_{LIML}-\beta_0\right)\right\|_2\left\|\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right\|_2\left\|\sqrt{N}\left(\hat{\beta}_{TSLS}-\beta_0\right)\right\|_2+o_p(1).\end{aligned}$$

This can be rewritten to give

$$\left\|\sqrt{N}\left(\hat{\beta}_{LIML}-\beta_0\right)\right\|_2\leqslant 2\cdot\frac{\left\|\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right\|_2}{\lambda_m\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)}\left\|\sqrt{N}\left(\hat{\beta}_{TSLS}-\beta_0\right)\right\|_2+o_p(1)$$

Notice that as N tends to infinity $\frac{\left\|\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right\|_2}{\lambda_m\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)}\to\frac{\left\|\Pi_{22}'H\Pi_{22}\right\|_2}{\lambda_m\left(\Pi_{22}'H\Pi_{22}\right)}$ which is finite and well defined and $\left\|\sqrt{N}\left(\hat{\beta}_{TSLS}-\beta_0\right)\right\|_2=O_p(1)$ so that $\left\|\sqrt{N}\left(\hat{\beta}_{LIML}-\beta_0\right)\right\|_2=O_p(1)$. Uniform tightness of $b_N=\sqrt{N}\left(\hat{\beta}_{LIML}-\beta_0\right)$ follows from the fact that $b_N=\sqrt{N}\left(\hat{\beta}_{LIML}-\beta_0\right)$ is a weakly convergent sequence.

We regard $N\times L_N\left(\beta_0+N^{-1/2}b\right)$ as a function of b on the set $B\subset R^p$, where B is compact. Multiplying N on both sides of (B.1), we obtain that

$$\begin{aligned}&N\times L_N\left(\beta_0+N^{-1/2}b_N\right)\\ =\;&\frac{N\hat{\pi}_{12}'\hat{H}^{1/2}M\hat{H}_{\hat{H}^{1/2}\hat{\Pi}_{22}}\hat{H}^{1/2}\hat{\pi}_{12}+\left(b_N-N^{1/2}\left(\hat{\beta}_{TSLS}-\beta_0\right)\right)'\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\left(b_N-N^{1/2}\left(\hat{\beta}_{TSLS}-\beta_0\right)\right)}{\left(1,-\beta_0'-N^{-1/2}b_N'\right)\hat{\Omega}\left(1,-\beta_0'-N^{-1/2}b_N'\right)'}\end{aligned}$$

where $\sqrt{N}\left(\hat{\beta}_{TSLS}-\beta_0\right)\to^D A(F_T)\,N\left(0,I_{(k_1+k_2)(p+1)}\right)$ by Theorem 1.

Let $\hat{C}$ be a matrix such that $\hat{C}\hat{C}'=M_{\hat{H}^{1/2}\hat{\Pi}_{22}}$ and $\hat{C}'\hat{C}=I_{k_2}$. By construction, $\hat{C}'\hat{H}^{1/2}\hat{\Pi}_{22}=0$, so

$$\hat{C}'\hat{H}^{1/2}\hat{\pi}_{12}=\hat{C}'\hat{H}^{1/2}\left(\hat{\pi}_{12}-\hat{\Pi}_{22}\beta_0\right).$$

Then

$$N\hat{\pi}_{12}'\hat{H}^{1/2}M\hat{H}_{\hat{H}^{1/2}\hat{\Pi}_{22}}\hat{H}^{1/2}\hat{\pi}_{12} = \left(\hat{C}'\hat{H}^{1/2}\sqrt{N}\left(\hat{\pi}_{12} - \hat{\Pi}_{22}\beta_0\right)\right)'\hat{C}'\hat{H}^{1/2}\sqrt{N}\left(\hat{\pi}_{12} - \hat{\Pi}_{22}\beta_0\right)$$

of which the probability limit (denoted by W) is uniformly bounded in $b \in B$ by noting , and the fact that $\hat{C} \to C$in which $CC' = M_{H^{1/2}\Pi_{22}}$ and $C'C = I_{k_2}$.

Then we obtain that

$$N \times L_N\left(\beta_0 + N^{-1/2}b_N\right) \to^D L(b) = \frac{W + (b - b_1)'\Pi_{22}'H(F_T)\Pi_{22}(b - b_1)}{(1, -\beta_0')\,\Omega\,(1, -\beta_0')'},$$

where $b_1 \equiv A(F_T)N\left(0, I_{(h_1+h_2+h_3+h_4)(p+1)}\right)$. Notice that $L(b)$ is continuous and is minimised at $b = A(F_T)N\left(0, I_{(h_1+h_2+h_3+h_4)(p+1)}\right)$. The proof is then complete. □

Proof of Theorem 2.

The first part follows from the fact that conditional on $\mathcal{F}$, $\hat{H} \to H(F_T)$ a.s., and

$$\begin{pmatrix} \hat{\pi}_{11} & \hat{\Pi}_{21} \\ \hat{\pi}_{12} & \hat{\Pi}_{22} \end{pmatrix} \to \begin{pmatrix} \pi_{11} & \Pi_{21} \\ \pi_{12} & \Pi_{22} \end{pmatrix} + \begin{pmatrix} \Delta_{11}(F_T) & \Delta_{21}(F_T) \\ \Delta_{12}(F_T) & \Delta_{22}(F_T) \end{pmatrix} \quad a.s.$$

To prove the second part notice that the LIML estimator minimizes

$$\frac{(1, -\beta')\left(\hat{\pi}_{12}, \hat{\Pi}_{22}\right)'\hat{H}\left(\hat{\pi}_{12}, \hat{\Pi}_{22}\right)(1, -\beta')'}{(1, -\beta')\,\hat{\Omega}\,(1, -\beta')'} \tag{B.2}$$

as a function of β on the compact set $B \subset R^p$. Notice that (B.2) is continuous for every β in this compact set B so that the continuity is also uniform.

By the Lemma B.1 and the proof of Lemma 2, we know (B.2) converges a.s. and uniformly to

$$\frac{(1, -\beta')(\pi_{12} + \Delta_{12}(F_T), \Pi_{22} + \Delta_{22}(F_T))'H(F_T)(\pi_{12} + \Delta_{12}(F_T), \Pi_{22} + \Delta_{22}(F_T))(1, -\beta')'}{(1, -\beta')\,\Omega(F_T)\,(1, -\beta')'}. \tag{B.3}$$

(B.3) has a unique minimizer. Thus, the LIML estimator converges to the minimum of (B.3).

Notice also that (B.3) is equal to zero when $H(F_T)(\pi_{12} + \Delta_{12}(F_T), \Pi_{22} + \Delta_{22}(F_T))(1, -\beta') = 0$. Since $\pi_{12} = \Pi_{22}\beta_0$, one has

$$H(F_T)(\Pi_{22} + \Delta_{22}(F_T))\beta = H(F_T)(\Pi_{22}\beta_0 + \Delta_{12}(F_T)).$$

Solving for β completes the proof. □

Proof of Theorem 3.

Conditional almost sure convergence can be easily proved, So, using Lemma 1, it follows that $\hat{\alpha}_0 = \hat{\pi}_{11} - \hat{\Pi}_{21}\hat{\beta} \to \pi_{11} - \Pi_{21}\beta = \alpha_0$ a.s. conditional on $\mathcal{F}$.

Since the panel LIML and TSLS estimators of the vector β_0 are asymptotically equivalent we prove this result for the TSLS estimator only. Notice that

$$\begin{aligned}
\hat{\alpha}_{TSLS} &= \left(\left(\hat{\pi}_{11}, \hat{\Pi}_{21}\right) - (\pi_{11}, \Pi_{21})\right)\begin{pmatrix} 1 \\ -\hat{\beta}_{TSLS} \end{pmatrix} + \pi_{11} - \Pi_{21}\hat{\beta}_{TSLS} \\
&= \left(\left(\hat{\pi}_{11}, \hat{\Pi}_{21}\right) - (\pi_{11}, \Pi_{21})\right)\begin{pmatrix} 1 \\ -\hat{\beta}_{TSLS} \end{pmatrix} + \pi_{11} - \Pi_{21}\left(\hat{\beta}_{TSLS} - \beta_0\right) - \Pi_{21}\beta_0 \\
&= (I_{k_1}, 0)\left(\hat{\Pi} - \Pi\right)\begin{pmatrix} 1 \\ -\hat{\beta}_{TSLS} \end{pmatrix} - \Pi_{21}\left(\hat{\beta}_{TSLS} - \beta_0\right) + \alpha_0 \\
&= (I_{k_1}, 0)\left(\hat{\Pi} - \Pi\right)\begin{pmatrix} 1 \\ -\hat{\beta}_{TSLS} \end{pmatrix} - \Pi_{21}\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)^{-1}\hat{\Pi}_{22}'\hat{H}\left(\hat{\pi}_{12} - \hat{\Pi}_{22}\beta_0\right) + \alpha_0
\end{aligned}$$

$$= \left(I_{k_1}, \underset{(k_1\times k_2)}{0}\right)\left(\hat{\Pi}-\Pi\right)\begin{pmatrix}1\\-\hat{\beta}_{TSLS}\end{pmatrix} - \Pi_{21}\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)^{-1}\hat{\Pi}_{22}'\hat{H}\left(0, I_{k_2}\right)\left(\hat{\Pi}-\Pi\right)\begin{pmatrix}1\\-\beta_0\end{pmatrix} + \alpha_0.$$

Thus

$$\begin{aligned}
\sqrt{N}\left(\hat{\alpha}_0-\alpha_0\right) &= \left(I_{k_1},0\right)\sqrt{N}\left(\hat{\Pi}-\Pi\right)\begin{pmatrix}1\\-\hat{\beta}_{TSLS}\end{pmatrix}\\
&\quad -\Pi_{21}\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)^{-1}\hat{\Pi}_{22}'\hat{H}\left(0,I_{k_2}\right)\sqrt{N}\left(\hat{\Pi}-\Pi\right)\begin{pmatrix}1\\-\beta_0\end{pmatrix}\\
&= \left(I_{k_1},0\right)\sqrt{N}\left(\hat{\Pi}-\Pi\right)\begin{pmatrix}0\\\beta_0-\hat{\beta}_{TSLS}\end{pmatrix}\\
&\quad +\left(\left(I_{k_1},0\right)-\Pi_{21}\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)^{-1}\hat{\Pi}_{22}'\hat{H}\left(0,I_{k_2}\right)\right)\sqrt{N}\left(\hat{\Pi}-\Pi\right)\begin{pmatrix}1\\-\beta_0\end{pmatrix}.
\end{aligned}$$

Notice that conditional on $\mathcal{F}$, $\left(I_{k_1},0\right)\sqrt{N}\left(\hat{\Pi}-\Pi\right)\begin{pmatrix}0\\\beta_0-\hat{\beta}_{TSLS}\end{pmatrix}\to 0$ a.s. and

$$\begin{aligned}
&\left(I_{k_1},0\right)-\Pi_{21}\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)^{-1}\hat{\Pi}_{22}'\hat{H}\left(0,I_{k_2}\right)\\
\to\ &\left(I_{k_1},0\right)-\Pi_{21}\left({\Pi_{22}}'H\left(F_T\right)\Pi_{22}\right)^{-1}{\Pi_{22}}'H\left(F_T\right)\left(0,I_{k_2}\right)\ a.s.
\end{aligned}$$

The asymptotic distribution follows from

$$\begin{aligned}
&\left(\left(I_{k_1},0\right)-\Pi_{21}\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)^{-1}\hat{\Pi}_{22}'\hat{H}\left(0,I_{k_2}\right)\right)\sqrt{N}\left(\hat{\Pi}-\Pi\right)\begin{pmatrix}1\\-\beta_0\end{pmatrix}\\
=\ & vec\left[\left(\left(I_{k_1},0\right)-\Pi_{21}\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)^{-1}\hat{\Pi}_{22}'\hat{H}\left(0,I_{k_2}\right)\right)\sqrt{N}\left(\hat{\Pi}-\Pi\right)\begin{pmatrix}1\\-\beta_0\end{pmatrix}\right]\\
=\ & \left(\left(1,-{\beta_0}'\right)\otimes\left(\left(I_{k_1},0\right)-\Pi_{21}\left(\hat{\Pi}_{22}'\hat{H}\hat{\Pi}_{22}\right)^{-1}\hat{\Pi}_{22}'\hat{H}\left(0,I_{k_2}\right)\right)\right)\left(\sqrt{N}vec\left[\hat{\Pi}-\Pi\right]\right).
\end{aligned}$$

Theorem 3 follows easily. □

Proof of Theorem 4. In the proofs of Lemma 2 and Theorem 2, we have shown

$$\begin{aligned}
\hat{\alpha}_{TSLS} &\to \pi_{11}+\Delta_{11}\left(F_T\right)-\left(\Pi_{21}+\Delta_{21}\left(F_T\right)\right)\left(\beta_0+b\left(F_T\right)\right)\quad a.s.,\\
\hat{\alpha}_{LIML} &\to \pi_{11}+\Delta_{11}\left(F_T\right)-\left(\Pi_{21}+\Delta_{21}\left(F_T\right)\right)\left(\beta_0+b\left(F_T\right)\right)\quad a.s.,
\end{aligned}$$

conditional on $\mathcal{F}$. Thus, the results follow immediately. □

Proof of Lemma 3

For simplicity let $\gamma=\gamma\left(F_T\right)$ and notice that

$$\begin{aligned}
\hat{\Theta} &= \frac{1}{N}\sum_{i=1}^{N} vec\left[\left(z_i-\bar{z}\right)'\left(e_i-F_T\gamma\right)\right]vec\left[\left(z_i-\bar{z}\right)'\left(e_i-F_T\gamma\right)\right]'\\
&-\frac{1}{N}\sum_{i=1}^{N} vec\left[\left(z_i-\bar{z}\right)'\left(e_i-F_T\gamma\right)\right]vec\left[\left(z_i-\bar{z}\right)'\left(\left(\hat{\pi}_{10},\hat{\Pi}_{20}\right)-\left(\pi_{10},\Pi_{20}\right)-F_T\gamma\right)\right]'\\
&-\frac{1}{N}\sum_{i=1}^{N} vec\left[\left(z_i-\bar{z}\right)'\left(e_i-F_T\gamma\right)\right]vec\left[\left(z_i-\bar{z}\right)'\left(z_i\left(\hat{\Pi}-\Pi\right)\right)\right]'\\
&-\frac{1}{N}\sum_{i=1}^{N} vec\left[\left(z_i-\bar{z}\right)'\left(\left(\hat{\pi}_{10},\hat{\Pi}_{20}\right)-\left(\pi_{10},\Pi_{20}\right)-F_T\gamma\right)\right]vec\left[\left(z_i-\bar{z}\right)'\left(e_i-F_T\gamma\right)\right]'\\
&+\frac{1}{N}\sum_{i=1}^{N} vec\left[\left(z_i-\bar{z}\right)'\left(\left(\hat{\pi}_{10},\hat{\Pi}_{20}\right)-\left(\pi_{10},\Pi_{20}\right)-F_T\gamma\right)\right]vec\left[\left(z_i-\bar{z}\right)'\left(\left(\hat{\pi}_{10},\hat{\Pi}_{20}\right)-\left(\pi_{10},\Pi_{20}\right)-F_T\gamma\right)\right]'
\end{aligned}$$

$$+\frac{1}{N}\sum_{i=1}^{N} vec\left[(z_i-\bar{z})'\left(\left(\hat{\pi}_{10},\hat{\Pi}_{20}\right)-(\pi_{10},\Pi_{20})-F_T\gamma\right)\right] vec\left[(z_i-\bar{z})' z_i\left(\hat{\Pi}-\Pi\right)\right]'$$
$$-\frac{1}{N}\sum_{i=1}^{N} vec\left[(z_i-\bar{z})' z_i\left(\hat{\Pi}-\Pi\right)\right] vec\left[(z_i-\bar{z})'(e_i-F_T\gamma)\right]'$$
$$+\frac{1}{N}\sum_{i=1}^{N} vec\left[(z_i-\bar{z})' z_i\left(\hat{\Pi}-\Pi\right)\right] vec\left[(z_i-\bar{z})'\left(\left(\hat{\pi}_{10},\hat{\Pi}_{20}\right)-(\pi_{10},\Pi_{20})-F_T\gamma\right)\right]'$$
$$+\frac{1}{N}\sum_{i=1}^{N} vec\left[(z_i-\bar{z})' z_i\left(\hat{\Pi}-\Pi\right)\right] vec\left[(z_i-\bar{z})' z_i\left(\hat{\Pi}-\Pi\right)\right]'.$$

Apart from the first, all terms vanish because they involve quantities which go to zero a.s. conditional on $\mathcal{F}$ given Assumpton 1.i-vi. We therefore focus on the first term.

Write

$$\frac{1}{N}\sum_{i=1}^{N} vec\left[(z_i-\bar{z})'(e_i-F_T\gamma)\right] vec\left[(z_i-\bar{z})'(e_i-F_T\gamma)\right]'$$

as

$$\frac{1}{N}\sum_{i=1}^{N} vec\left[\left(z_i-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right] vec\left[\left(z_i-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right]' \tag{B.4}$$
$$-\frac{1}{N}\sum_{i=1}^{N} vec\left[\left(z_i-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right] vec\left[\left(\bar{z}-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right]' \tag{B.5}$$
$$-\frac{1}{N}\sum_{i=1}^{N} vec\left[\left(\bar{z}-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right] vec\left[\left(z_i-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right]' \tag{B.6}$$
$$+\frac{1}{N}\sum_{i=1}^{N} vec\left[\left(\bar{z}-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right] vec\left[\left(\bar{z}-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right]' \tag{B.7}$$

The first term (B.4) converges a.s. to $\Theta(F_T)$ since each term involves squares of the components of $e_i-F_T\gamma$ and squares of the components of z_i. The expectations of the first quantities are a.s. bounded by Assumption 1.i and 1.iv, while those of the second components are a.s. bounded by Assumption 1.ii, iii and v:

$$E\left[\left\|vec\left[\left(z_i-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right] vec\left[\left(z_i-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right]'\right\|_2 |\mathcal{F}\right]$$
$$=E\left[\left\|\left(z_i-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right\|_2^2 |\mathcal{F}\right]$$
$$\leqslant E\left[\left\|z_i-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right\|_2^2 |\mathcal{F}\right] E\left[\left\|e_i-F_T\gamma\right\|_2^2 |\mathcal{F}\right] \text{ Subadditivity and conditional independence}$$

where the last quantities are bounded a.s. by Assumption 1.

The term (B.5) converges a.s. to zero. To see this, notice that

$$vec\left[\left(\bar{z}-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right]=\left(I_{p+1}\otimes\left(\bar{z}-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)\right) vec\left[e_i-F_T\gamma\right]$$

So that (B.5) can be written as

$$\left(\frac{1}{N}\sum_{i=1}^{N} vec\left[\left(z_i-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'(e_i-F_T\gamma)\right] vec\left[e_i-F_T\gamma\right]'\right)\left(I_{p+1}\otimes\left(\bar{z}-\frac{1}{N}\sum_{j=1}^{N}E\left[z_j|\mathcal{F}\right]\right)'\right).$$

The term $\bar{z} - \frac{1}{N}\sum_{j=1}^{N} E[z_j|\mathcal{F}] \to 0$ a.s.. Moreover, each term in the sum has zero mean and each term involves products of components of $(e_i - F_T\gamma)(e_i - F_T\gamma)'$ and components of $z_i - \frac{1}{N}\sum_{j=1}^{N} E[z_j|\mathcal{F}]$. The moments of these products are a.s. bounded by Assumption 1.i, iv, ii, iii and v.:

$$
\begin{aligned}
& E\left[\left\| vec\left[\left(z_i - \frac{1}{N}\sum_{j=1}^{N} E[z_j|\mathcal{F}]\right)'(e_i - F_T\gamma)\right] vec[e_i - F_T\gamma]'\right\|_2 |\mathcal{F}\right] \\
& \leqslant E\left[\left\|\left(z_i - \frac{1}{N}\sum_{j=1}^{N} E[z_j|\mathcal{F}]\right)'(e_i - F_T\gamma)\right\|_2 \|e_i - F_T\gamma\|_2 |\mathcal{F}\right] \text{ Cauchy - Schwarz inequality} \\
& \leqslant E\left[\left\| z_i - \frac{1}{N}\sum_{j=1}^{N} E[z_j|\mathcal{F}]\right\|_2 |\mathcal{F}\right] E[\|e_i - F_T\gamma\|_2 |\mathcal{F}]^2 \text{ Subadditivity and conditional independence}
\end{aligned}
$$

The two expectations above are bounded uniformly a.s. by Assumption 1.

(B.6) is the transpose of (B.5).

(B.7) can be written as

$$
\left(I_{p+1} \otimes \left(\bar{z} - \frac{1}{N}\sum_{j=1}^{N} E[z_j|\mathcal{F}]\right)\right)\left(\frac{1}{N}\sum_{i=1}^{N} vec[e_i - F_T\gamma]\, vec[e_i - F_T\gamma]'\right)\left(I_{p+1} \otimes \left(\bar{z} - \frac{1}{N}\sum_{j=1}^{N} E[z_j|\mathcal{F}]\right)\right)'.
$$

Notice that $\bar{z} - \frac{1}{N}\sum_{j=1}^{N} E[z_j|\mathcal{F}] \to 0$ a.s. and the term in the middle is a.s. bounded since it involves only terms of $(e_i - F_T\gamma)(e_i - F_T\gamma)'$ by Assumption 1.i and iv. Thus, (B.7) converges to zero a.s. conditional on $\mathcal{F}$. $\square$

Appendix C

This file provides the proofs of Theorem A.1, Theorem A.2 and Lemma B.1. The proofs of Theorems A.1 and A.2 follow the classical proofs and are reported here just for the sake of completeness.

Proposition C.1. (Conditional generalized Kolmogorov inequality, Rao (2009, Theorem 4, p. 449)) Assume that $\{X_k : 1 \leqslant k \leqslant n\}$ is a set of $\mathcal{F}$-independent random variables with $E[|X_k|^r|\mathcal{F}] < \infty$ for each k and some $r \geqslant 1$ where $E[\cdot|\mathcal{F}]$ denotes the conditional expectation given the sub-σ-algebra $\mathcal{F}$. For any $\mathcal{F}$-measurable random variable $\varepsilon > 0$ a.s., let $S_k = \sum_{n=1}^{k} X_n$ be the partial sum and let event

$$
D = \left(\max_{1\leqslant k\leqslant n} |S_k - E[S_k|\mathcal{F}]| \geqslant \varepsilon\right),
$$

then we have

$$
\varepsilon^r P(D|\mathcal{F}) \leqslant E[|S_k - E(S_k|\mathcal{F})|^r \cdot I_D|\mathcal{F}] \leqslant E[|S_k - E[S_k|\mathcal{F}]|^r|\mathcal{F}] \qquad a.s.,
$$

where $I_D = 1$ when D holds and $I_D = 0$ otherwise.

Proposition C.2. (Conditional Borel-Cantelli lemma, Rao (2009, Theorem 1, pp.444-445); Majerek et al. (2005, Theorem 3.1, Lemmas 3.2 and 3.3, pp. 149-151)) Suppose that $(\Omega, \mathcal{A}, P)$ is a probability space and $\mathcal{F}$ is a sub-σ-algebra of $\mathcal{A}$. We have the following results:

1. Let $\{A_n, n \geqslant 1\}$ be a sequence of $\mathcal{F}$-independent events such that $\sum_{n=1}^{\infty} P(A_n) < \infty$. Then $\sum_{n=1}^{\infty} P(A_n|\mathcal{F}) < \infty$ a.s..
2. Let $\{A_n, n \geqslant 1\}$ be a sequence of $\mathcal{F}$-independent events and $A = \left\{\omega : \sum_{n=1}^{\infty} E[I_{A_n}|\mathcal{F}] = \infty\right\}$ with $P(A) < 1$. Then only finitely many events from the event sequence $\{A_n \cap A, n \geqslant 1\}$ hold with probability one.

3. Let $\{A_n, n \geqslant 1\}$ be a sequence of $\mathcal{F}$-independent events and let $A = \left\{\omega : \sum_{n=1}^{\infty} E\left[I_{A_n}|\mathcal{F}\right] = \infty\right\}$. Then it holds that $P\left(\limsup_{n\to\infty} A_n\right) = P(A)$.

Proposition C.3. (Kronecker's lemma, Chow and Teicher (1997, Lemma 2, pp. 114-115)) If $\{a_n\}$ and $\{b_n\}$ are sequences of real numbers where $0 < b_n \uparrow \infty$ and $\sum_{i=1}^{\infty} a_i/b_i$ converges, we have $\left(\sum_{i=1}^{n} a_i\right)/b_n \to 0$.

Proposition C.4. (Bound on Characteristic Function, Chow and Teicher (1997, Lemma 1, p. 295)) For any $t \in (-\infty, \infty)$ and arbitrary nonnegative integer n

$$e^{it} - \sum_{j=0}^{n} \frac{(it)^k}{k!} = \frac{(it)^{n+1}}{n!} \int_0^1 e^{itu}(1-u)^n du = i^{n+1} \int_0^t dt_{n+1} \int_0^{t_{n+1}} dt_n \cdots \int_0^{t_2} e^{it_1} dt_1.$$

Moreover, for each $0 \leqslant \delta \leqslant 1$,

$$\left| e^{it} - \sum_{j=0}^{n} \frac{(it)^k}{k!} \right| \leqslant \frac{2^{1-\delta}|t|^{n+\delta}}{(1+\delta)(2+\delta)\cdots(n+\delta)}.$$

Proposition C.5. (Conditional expectation of a product lemma, Chow and Teicher (1997, Corollary 5, p. 234)) Let $X_j \in \mathcal{L}_1$ be the space of all measurable function with finite mean and $\mathcal{B}^\infty$ be the class of Borel subsets of $R^\infty = R \times R \times \cdots$. If the random variables $\{X_n, n \geqslant 1\}$ are conditionally independent given the σ-algebra $\mathcal{F}$ of events, then there exists a regular conditional distribution P^ω for $X = (X_1, X_2, \ldots)$ given $\mathcal{F}$ such that for each $\omega \in \Omega$ the coordinate random variable sequence $\{\zeta_n, n \geqslant 1\}$ of the probability space $(R^\infty, \mathcal{B}^\infty, P^\omega)$ are independent. Moreover, if $X_j \in \mathcal{L}_1$ for every $1 \leqslant j \leqslant n$ and $E\left[X_1 X_1 \cdots X_n|\mathcal{F}\right]$ exists, then when $n \geqslant 2$ it holds that

$$E\left[X_1 X_1 \cdots X_n|\mathcal{F}\right] = \prod_{i=1}^{n} E\left[X_i|\mathcal{F}\right] \qquad a.s..$$

Proposition C.6. (Jensen's inequality for conditional expectations, Chow and Teicher (1997, Theorem 4, p. 217)) Let $g \in \mathbb{R} \to \mathbb{R}$ be convex. Then for any Y such that $g(Y)$ is integrable,

$$E\left[g(Y)|\mathcal{F}\right] \geqslant g\left(E\left[Y|\mathcal{F}\right]\right) \qquad a.s..$$

Theorem C.1. (Conditional Markov strong law of large numbers) Let $\{Z_i : i \geqslant 1\}$ be a sequence of $\mathcal{F}$-independent random variables with conditional means $E\left[Z_i|\mathcal{F}\right]$ for $i = 1, 2, \ldots$. If for some scalar $0 < \delta \leqslant 1$, $\sum_{i=1}^{\infty} \frac{1}{i^{1+\delta}} E\left[\left|Z_i - E\left[Z_i|\mathcal{F}\right]\right|^{1+\delta}|\mathcal{F}\right] < \infty$ a.s., then conditional on $\mathcal{F}$, $\frac{1}{n}\sum_{i=1}^{n}\left(Z_i - E\left[Z_i|\mathcal{F}\right]\right) \to 0$ a.s..

Proof of Theorem C.1:

The proof of this result follows closely the proof of classical (unconditional) case given in Chung (1974, pp. 130-132). Let $\mu_i = E\left[Z_i|\mathcal{F}\right]$. First of all, we define for each i,

$$Y_i = \begin{cases} Z_i - \mu_i & |Z_i - \mu_i| \leqslant i, \\ 0 & |Z_i - \mu_i| > i. \end{cases}$$

Then,

$$\sum_{i=1}^{\infty} E\left[i^{-2}Y_i^2|\mathcal{F}\right] = \sum_{i=1}^{\infty} E\left[i^{-2}|Z_i - \mu_i|^2 \cdot 1\left\{|Z_i - \mu_i| \leqslant i\right\}|\mathcal{F}\right] \qquad a.s..$$

Notice that for any real number $|y| \leqslant i$ and $0 < \delta \leqslant 1$, $y^2/i^2 \leqslant y^{1+\delta}/i^{1+\delta}$, so

$$\begin{aligned}\sum_{i=1}^{\infty} Var\left[i^{-1}Y_i|\mathcal{F}\right] &\leqslant \sum_{i=1}^{\infty} E\left[i^{-2}Y_i^2|\mathcal{F}\right] \leqslant \sum_{i=1}^{\infty} E\left[i^{-1-\delta}|Z_i - \mu_i|^{1+\delta} \cdot 1\left\{|Z_i - \mu_i| \leqslant i\right\}|\mathcal{F}\right] \\ &\leqslant \sum_{i=1}^{\infty} E\left[i^{-1-\delta}|Z_i - \mu_i|^{1+\delta}|\mathcal{F}\right] < \infty \quad a.s..\end{aligned}$$

Moreover, $\{Y_i\}$ is also a $\mathcal{F}$-independent random sequence by construction. Therefore, we can apply the generalized Kolmogorov inequality. For any $\mathcal{F}$-measurable $m \geqslant 1$

$$P\left(\max_{k\leqslant j\leqslant l}\left|\sum_{i=k}^{j}\left(i^{-1}Y_i - E\left[i^{-1}Y_i|\mathcal{F}\right]\right)\right| \leqslant m^{-1}|\mathcal{F}\right) \geqslant 1 - m^2\sum_{i=k}^{l} Var\left[i^{-1}Y_i|\mathcal{F}\right] \qquad a.s..$$

Since $\sum_{i=1}^{\infty} Var\left[i^{-1}Y_i|\mathcal{F}\right]$ is convergent a.s., we must have

$$\lim_{k\to\infty}\lim_{l\to\infty} P\left(\max_{k\leqslant j\leqslant l}\left|\sum_{i=k}^{j}\left(i^{-1}Y_i - E\left[i^{-1}Y_i|\mathcal{F}\right]\right)\right| \leqslant m^{-1}|\mathcal{F}\right) = 1 \qquad a.s..$$

Thus, the tail of $\sum_{i=1}^{\infty}\left(i^{-1}Y_i - E\left[i^{-1}Y_i|\mathcal{F}\right]\right)$ converges to zero a.s. conditional on $\mathcal{F}$ and consequently $\sum_{i=1}^{\infty}\left(i^{-1}Y_i - E\left[i^{-1}Y_i|\mathcal{F}\right]\right)$ converges a.s. conditional on $\mathcal{F}$. Second, we prove that $\sum_{i=1}^{\infty} E\left[i^{-1}Y_i|\mathcal{F}\right]$ converges a.s.. Notice that for each real number $|y| > i$ and $0 < \delta \leqslant 1$, $|y|/i \leqslant y^{1+\delta}/i^{1+\delta}$, so

$$\begin{aligned}
\sum_{i=1}^{\infty} E\left[i^{-1}Y_i|\mathcal{F}\right] &\leqslant \sum_{i=1}^{\infty} E\left[i^{-1-\delta}|Z_i-\mu_i|^{1+\delta}\cdot 1\{|Z_i-\mu_i|<i\}|\mathcal{F}\right] \\
&\leqslant \sum_{i=1}^{\infty} E\left[i^{-1-\delta}|Z_i-\mu_i|^{1+\delta}|\mathcal{F}\right] < \infty \qquad a.s..
\end{aligned}$$

Since $\sum_{i=1}^{\infty}\left(i^{-1}Y_i - E\left[i^{-1}Y_i|\mathcal{F}\right]\right)$ and $\sum_{i=1}^{\infty} E\left[i^{-1}Y_i|\mathcal{F}\right]$ converge a.s. conditional on $\mathcal{F}$, $\sum_{i=1}^{\infty} i^{-1}Y_i$ converges a.s. conditional on $\mathcal{F}$, too. It remains to show that a.s. convergence of $\sum_{i=1}^{\infty} i^{-1}Y_i$ conditional on $\mathcal{F}$ implies a.s. convergence of $\sum_{i=1}^{\infty} i^{-1}(Z_i-\mu_i)$ conditional on $\mathcal{F}$. It is easy to check that

$$\begin{aligned}
&\sum_{i=1}^{\infty} P\left(Y_i \neq Z_i-\mu_i|\mathcal{F}\right) = \sum_{i=1}^{\infty} E\left[1\{|Z_i-\mu_i|>i\}|\mathcal{F}\right] \\
&\leqslant \sum_{i=1}^{\infty} E\left[i^{-1-\delta}|Z_i-\mu_i|^{1+\delta}\cdot 1\{|Z_i-\mu_i|>i\}|\mathcal{F}\right] \leqslant \sum_{i=1}^{\infty} E\left[i^{-1-\delta}|Z_i-\mu_i|^{1+\delta}|\mathcal{F}\right] < \infty \qquad a.s..
\end{aligned}$$

If we define a sequence of events $\mathcal{E}_i = \{\omega : Y_i \neq Z_i - \mu_i\}$ for $i \geqslant 1$, we know that $\mathcal{E}_i$ is a sequence of $\mathcal{F}$-independent events. Due to part (3) of the conditional Borel-Cantelli lemma, the conditional probability of the event that $Y_i \neq Z_i-\mu_i$ holds infinitely often is zero conditional on $\mathcal{F}$. Thus, conditional on $\mathcal{F}$, $\sum_{i=1}^{\infty} i^{-1}(Z_i-\mu_i)$ converges a.s.. The application of Kronecker's lemma to $\sum_{i=1}^{\infty} i^{-1}(Z_i-\mu_i)$ for each $\omega\in\Omega$, implies that $\bar{Z}_n - \bar{\mu}_n \to 0$ a.s. conditional on $\mathcal{F}$. ■

Theorem C.2 (**Conditional Lindeberg central limit theorem**)**.** Let $\{Z_i : i \geqslant 1\}$ be a sequence of $\mathcal{F}$-independent random variables with conditional means $E[Z_i|\mathcal{F}]$ and conditional variances $\sigma_i^2 = E\left[(Z_i - E[Z_i|\mathcal{F}])^2|\mathcal{F}\right]$ for $i = 1, 2, ...$ If there is η $\mathcal{F}$-measurable such that $\bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 > \eta > 0$ a.s. and the following conditional Lindeberg condition holds

$$\lim_{n\to\infty}\frac{1}{n\bar{\sigma}_n^2}\sum_{i=1}^{n} E\left[(Z_i - E[Z_i|\mathcal{F}])^2\cdot 1\left\{|Z_i - E[Z_i|\mathcal{F}]| > \sqrt{n}\bar{\sigma}_n\varepsilon\right\}|\mathcal{F}\right] = 0 \text{ a.s.}$$

for any $\mathcal{F}$-measurable $\varepsilon > 0$, then conditional on $\mathcal{F}$, $\frac{1}{\bar{\sigma}_n\sqrt{n}}\sum_{i=1}^{n}(Z_i - E[Z_i|\mathcal{F}]) \to^D N(0,1)$ a.s.. Moreover, $\frac{1}{\bar{\sigma}_n\sqrt{n}}\sum_{i=1}^{n}(Z_i - E[Z_i|\mathcal{F}]) \to^D N(0,1)$ ($\mathcal{F}$-stably).

Proof of Theorem C.2:

The theorem is proved similarly to the classical version (see, for example, Chow and Teicher (1997, pp. 314-315)) with appropriate modifications. Let $X_k = \frac{Z_k - \mu_k}{\sqrt{n}\bar{\sigma}_n}$, where $\mu_i = E\left[Z_i|\mathcal{F}\right]$. For any fixed real number $t \in R$, write $Y_k(t) = e^{itX_k} - 1 - itX_k + \frac{t^2 X_k^2}{2}$ and $y_k(t) = e^{-E\left[X_k^2|\mathcal{F}\right]t^2/2} - 1 + \frac{E\left[X_k^2|\mathcal{F}\right]t^2}{2}$. Using the bound on the characteristic function summarized above with $\delta = 1$ and $n = 1, 2$, one obtains

$$|Y_k(t)| \leqslant \left|e^{itX_k} - 1 - itX_k\right| + \frac{t^2 X_k^2}{2} \leqslant \frac{t^2 X_k^2}{2} + \frac{t^2 X_k^2}{2} = t^2 X_k^2$$

and

$$|Y_k(t)| = \left|e^{itX_k} - 1 - itX_k + \frac{t^2 X_k^2}{2}\right| \leqslant \frac{\left|tX_k\right|^3}{6}.$$

So, $|Y_k(t)| \leqslant Min\left\{t^2 X_k^2, \frac{|tX_k|^3}{6}\right\}$. Moreover, $|y_k(t)| \leqslant \frac{\left(E\left[X_k^2|\mathcal{F}\right]\right)^2 t^4}{8}$. As a result, for arbitrary $\mathcal{F}$-measurable $\varepsilon > 0$, noticing that $E\left[X_k|\mathcal{F}\right] = 0$,

$$\begin{aligned}
&\left|E\left[e^{itX_k}|\mathcal{F}\right] - e^{-\sigma_k^2 t^2/2}\right| = \left|E\left[Y_k(t)|\mathcal{F}\right] - y_k(t)\right| \leqslant \left|E\left[Y_k(t)|\mathcal{F}\right]\right| + |y_k(t)| \\
&\leqslant E\left[t^2 X_k^2 \cdot 1\left\{|X_k| > \varepsilon\right\} + \left|tX_k\right|^3 \cdot 1\left\{|X_k| \leqslant \varepsilon\right\}|\mathcal{F}\right] + \frac{\left(E\left[X_k^2|\mathcal{F}\right]\right)^2 t^4}{8} \quad a.s..
\end{aligned}$$

Notice that for $0 < j \leqslant n$,

$$E\left[\left(\sum_{k=1}^{j} X_k\right)^2 |\mathcal{F}\right] = \sum_{k=1}^{j} E\left[X_k^2|\mathcal{F}\right] \leqslant \frac{1}{n\bar{\sigma}_n^2}\sum_{k=1}^{n} E\left[(Z_k - \mu_k)^2|\mathcal{F}\right] = 1 \quad a.s.,$$

because the X_k's are $\mathcal{F}$-independent. Notice that this implies that $\sum_{j=1}^{n} E\left[X_j^2|\mathcal{F}\right] = 1$ a.s.. Thus,

$$\begin{aligned}
&\left|E\left[\exp\left\{it\sum_{k=1}^{j} X_k + \frac{1}{2}t^2\sum_{k=1}^{j} E\left[X_k^2|\mathcal{F}\right]\right\}|\mathcal{F}\right] - E\left[\exp\left\{it\sum_{k=1}^{j-1} X_k + \frac{1}{2}t^2\sum_{k=1}^{j-1} E\left[X_k^2|\mathcal{F}\right]\right\}|\mathcal{F}\right]\right| \\
= \ &\left|E\left[\exp\left\{it\sum_{k=1}^{j-1} X_j + \frac{1}{2}t^2\sum_{k=1}^{j} E\left[X_k^2|\mathcal{F}\right]\right\}|\mathcal{F}\right] \cdot E\left[e^{itX_j} - e^{-E\left[X_j^2|\mathcal{F}\right]t^2/2}|\mathcal{F}\right]\right| \\
= \ &\left|E\left[\exp\left\{it\sum_{k=1}^{j-1} X_j\right\}|\mathcal{F}\right]\right| \exp\left\{\frac{1}{2}t^2\sum_{k=1}^{j} E\left[X_k^2|\mathcal{F}\right]\right\}\left|E\left[e^{itX_j} - e^{-E\left[X_j^2|\mathcal{F}\right]t^2/2}|\mathcal{F}\right]\right| \\
\leq \ &E\left[\left|\exp\left\{it\sum_{k=1}^{j-1} X_j\right\}\right||\mathcal{F}\right] \cdot \left|\exp\left\{\frac{1}{2}t^2\sum_{k=1}^{j} E\left[X_k^2|\mathcal{F}\right]\right\}\right| \cdot \left|E\left[e^{itX_j} - e^{-E\left[X_j^2|\mathcal{F}\right]t^2/2}|\mathcal{F}\right]\right| \quad a.s., \\
\leq \ &\left|E\left[\exp\left\{\frac{1}{2}t^2 \cdot 1\right\}|\mathcal{F}\right]\right| \cdot \left|E\left[e^{itX_j} - e^{-E\left[X_j^2|\mathcal{F}\right]t^2/2}|\mathcal{F}\right]\right| \quad a.s., \\
\leq \ &e^{t^2/2}\left|E\left[e^{itX_j} - e^{-E\left[X_j^2|\mathcal{F}\right]t^2/2}|\mathcal{F}\right]\right| \quad a.s., \\
\leq \ &e^{t^2/2}E\left[t^2 X_j^2 \cdot 1\left\{|X_j| > \varepsilon\right\} + |t|^3|X_j|^3 \cdot 1\left\{|X_j| \leqslant \varepsilon\right\}|\mathcal{F}\right] + \frac{\left(E\left[X_j^2|\mathcal{F}\right]\right)^2 t^4}{8} \quad a.s., \\
\leq \ &e^{t^2/2}E\left[t^2 X_j^2 \cdot 1\left\{|X_j| > \varepsilon\right\} + \varepsilon|t|^3|X_j|^2 \cdot 1\left\{|X_j| \leqslant \varepsilon\right\}|\mathcal{F}\right] + E\left[X_j^2|\mathcal{F}\right] t^4 \max_{1\leqslant i\leqslant n} E\left[X_i^2|\mathcal{F}\right] \quad a.s.,
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\left|E\left[\exp\left\{it\sum_{k=1}^{n} X_k\right\}|\mathcal{F}\right] - e^{-t^2/2}\right| \\
= \ &\left|e^{-t^2/2}\sum_{j=1}^{n} E\left[\exp\left(it\sum_{k=1}^{j} X_k + \sum_{k=1}^{j} E\left[X_k^2|\mathcal{F}\right]\frac{t^2}{2}\right) - \exp\left(it\sum_{k=1}^{j-1} X_k + \sum_{k=1}^{j-1} E\left[X_k^2|\mathcal{F}\right]\frac{t^2}{2}\right)|\mathcal{F}\right]\right| \\
\leq \ &e^{-t^2/2}\sum_{j=1}^{n}\left|E\left[\exp\left(it\sum_{k=1}^{j} X_k + \sum_{k=1}^{j} E\left[X_k^2|\mathcal{F}\right]\frac{t^2}{2}\right) - \exp\left(it\sum_{k=1}^{j-1} X_k + \sum_{k=1}^{j-1} E\left[X_k^2|\mathcal{F}\right]\frac{t^2}{2}\right)|\mathcal{F}\right]\right|
\end{aligned}$$

$$
\begin{aligned}
\leq\ & t^2 \sum_{j=1}^{n} E\left[X_j^2 \cdot 1\left\{|X_j| > \varepsilon\right\}|\mathcal{F}\right] + \varepsilon|t|^3 \sum_{j=1}^{n} E\left[X_j^2 \cdot 1\left\{|X_j| \leqslant \varepsilon\right\}|\mathcal{F}\right] \\
& + t^4 \max_{1\leqslant i\leqslant n} E\left[X_i^2|\mathcal{F}\right] \cdot \sum_{j=1}^{n} E\left[X_j^2|\mathcal{F}\right] \\
\leq\ & t^2 \sum_{j=1}^{n} E\left[X_j^2 \cdot 1\left\{|X_j| > \varepsilon\right\}|\mathcal{F}\right] + \varepsilon|t|^3 \sum_{j=1}^{n} E\left[X_j^2|\mathcal{F}\right] + t^4 \max_{1\leqslant i\leqslant n} E\left[X_i^2|\mathcal{F}\right] \cdot \sum_{j=1}^{n} E\left[X_j^2|\mathcal{F}\right] \\
\leq\ & t^2 \sum_{j=1}^{n} E\left[X_j^2 \cdot 1\left\{|X_j| > \varepsilon\right\}|\mathcal{F}\right] + \varepsilon|t|^3 + t^4 \max_{1\leqslant i\leqslant n} E\left[X_i^2|\mathcal{F}\right].
\end{aligned}
$$

The last term tends to zero almost surely when n is large since for arbitrarily small $\mathcal{F}$-measurable $\varepsilon > 0$ and $1 \leqslant i \leqslant n$, one has $X_i^2 \leqslant \varepsilon^2 + X_i^2 \cdot 1\{|X_i| > \varepsilon\}$ a.s. implying

$$
E\left[X_i^2|\mathcal{F}\right] \leqslant \varepsilon^2 + E\left[X_i^2 \cdot 1\left\{|X_i| > \varepsilon\right\}|\mathcal{F}\right] \qquad a.s.,
$$

which entails

$$
E\left[X_i^2|\mathcal{F}\right] \leqslant \varepsilon^2 + \sum_{i=1}^{n} E\left[X_i^2 \cdot 1\left\{|X_i| > \varepsilon\right\}|\mathcal{F}\right] \qquad a.s..
$$

Noticing that the right-hand side does not depend on i and the conditional Lindeberg condition implies that $\max_{1\leqslant i\leqslant n} E\left[X_i^2|\mathcal{F}\right] \leqslant \varepsilon^2 + \sum_{i=1}^{n} E\left[X_i^2 \cdot 1\left\{|X_i| > \varepsilon\right\}|\mathcal{F}\right] \to 0$ a.s. as $n \to \infty$ and $\varepsilon \to 0$. The terms in the middle converges to zero a.s. as $\varepsilon \to 0$ a.s., and the first term also converges to zero a.s. as $n \to \infty$ and $\varepsilon \to 0$ a.s. because of the conditional Lindeberg condition. Therefore, it follows that

$$
E\left[\exp\left\{it\sum_{k=1}^{n} X_k\right\}|\mathcal{F}\right] = E\left[\exp\left\{it\bar{\sigma}_n^{-1}\sqrt{n}\left(\bar{Z}_n - \bar{\mu}_n\right)\right\}|\mathcal{F}\right] \to e^{-t^2/2} \quad a.s.,
$$

when n tends to infinity. Hence, the Levy continuity theorem implies that conditionally on $\mathcal{F}$, $\sqrt{n}\bar{\sigma}_n^{-1}\left(\bar{Z}_n - \bar{\mu}_n\right) \to^D N(0,1)$. Since the right-hand side above does not depend on the conditioning sigma-algebra. The result must hold unconditionally. Due to equation (1.5) of Rényi (1963, p. 294) we observe that this convergence is also $\mathcal{F}$-stable. The result is proved. ■

As can be seen, Theorems A.1 and A.2 are in fact special cases of Theorems C.1 and C.2 respectively. Therefore, the proofs of Theorems A.1 and A.2 can be also easily derived following those of Theorems C.1 and C.2 and are given as follows.

Proof of Theorem A.1: Notice first that the c_r inequality (e.g. White (2001, Proposition 3.8, p. 35)) ensures that

$$
E\left(\left|Z_i - E\left(Z_i|\mathcal{F}\right)\right|^{1+\delta}|\mathcal{F}\right) \leqslant 2^\delta\left(E\left(\left|Z_i\right|^{1+\delta}|\mathcal{F}\right) + \left|E\left(Z_i|\mathcal{F}\right)\right|^{1+\delta}\right) \qquad a.s..
$$

By noting that $E\left(\left|Z_i\right|^{1+\delta}|\mathcal{F}\right) < \Delta$ a.s. and applying Jensen's inequality for conditional expectations, we have $\left|E\left(Z_i|\mathcal{F}\right)\right|^{1+\delta} \leqslant E\left(\left|Z_i\right|^{1+\delta}|\mathcal{F}\right) < \Delta$ a.s. so that $E\left(\left|Z_i - E\left(Z_i|\mathcal{F}\right)\right|^{1+\delta}|\mathcal{F}\right) \leqslant 2^{1+\delta}\cdot\Delta$ a.s.. Then, the moment condition given in Theorem C.1 can be easily verified and the proof is complete. ■

Proof of Theorem A.2: The proof follows immediately by verifying the conditional Lindeberg condition given in Theorem C.2. ■

Proof of Lemma B.1:

It should be noted that throughout this proof we set $0 < \delta < 1$ and that all results are conditional on $\mathcal{F}$. We frequently make use of Cramer-Wold device and Theorem A.1.

1. Let ζ_1 and ζ_2 be arbitrary $(h_1 + h_3) \times 1$ vectors. $\frac{1}{N}\sum_{i=1}^{N} \zeta_1{}' w_i{}' w_i \zeta_2$ is a sum of $\mathcal{F}$-independent random variables and each term satisfies the following inequality a.s.

$$
E\left[\left|\zeta_1{}' w_i{}' w_i \zeta_2\right|^{1+\delta}|\mathcal{F}\right] \leqslant \left\|\zeta_1\right\|_2{}^{1+\delta}\left\|\zeta_2\right\|_2{}^{1+\delta} E\left[\left\|w_i\right\|_2{}^{2+2\delta}|\mathcal{F}\right].
$$

The term on the right-hand side is a.s. uniformly bounded because of Assumption 1. Thus, from the use of the Cramer-Wold device and Theorem A.1 one has

$$\frac{1}{N}\sum_{i=1}^{N} w_i' w_i \to \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[w_i' w_i|\mathcal{F}\right] = W\left(F_T\right) \qquad a.s..$$

2. Let ζ_1 and ζ_2 be arbitrary $(h_2+h_4)\times 1$ vectors. $\frac{1}{N}\sum_{i=1}^{N}\zeta_1' x_i' x_i \zeta_2$ is a sum of $\mathcal{F}$-independent random variables and each term satisfies the following inequality a.s.

$$\begin{aligned} & E\left[\left|\zeta_1' x_i' x_i \zeta_2\right|^{1+\delta}|\mathcal{F}\right] \leqslant \left\|\zeta_1\right\|_2^{1+\delta}\left\|\zeta_2\right\|_2^{1+\delta} E\left[\left\|x_i\right\|_2^{2+2\delta}|\mathcal{F}\right] \\ \leq\ & 2^{1+2\delta}\left\|\zeta_1\right\|_2^{1+\delta}\left\|\zeta_2\right\|_2^{1+\delta}\left(E\left[\left\|v_i\right\|_2^{2+2\delta}|\mathcal{F}\right] + \left\|F_T\right\|_2^{2+2\delta} E\left[\left\|\Gamma_i\right\|_2^{2+2\delta}|\mathcal{F}\right]\right), \end{aligned}$$

where the last line follows the c_r inequality. The right-hand side is bounded uniformly due to Assumption 1.ii and v. Therefore,

$$\begin{aligned} & \frac{1}{N}\sum_{i=1}^{N} x_i' x_i = \frac{1}{N}\sum_{i=1}^{N} v_i' v_i + \frac{1}{N}\sum_{i=1}^{N}\Gamma_i' F_T' F_T \Gamma_i + \frac{1}{N}\sum_{i=1}^{N} v_i' F_T \Gamma_i + \frac{1}{N}\sum_{i=1}^{N}\Gamma_i' F_T' v_i \\ \to\ & V\left(F_T\right) + \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[\Gamma_i' F_T' F_T \Gamma_i|\mathcal{F}\right] \\ & + \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[v_i'|\mathcal{F}\right] F_T \Gamma\left(F_T\right) + \Gamma\left(F_T\right)' F_T' \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[v_i|\mathcal{F}\right] \\ =\ & X\left(F_T\right) \end{aligned}$$

a.s., which gives the result.

3. Let ζ_1 and ζ_2 be respectively arbitrary $(h_1+h_3)\times 1$ and $(h_2+h_4)\times 1$ vectors. $\frac{1}{N}\sum_{i=1}^{N}\zeta_1' w_i' x_i \zeta_2$ is a sum of $\mathcal{F}$-independent random variables and each term satisfies the following inequality a.s.

$$\begin{aligned} & E\left[\left|\zeta_1' w_i' x_i \zeta_2\right|^{1+\delta}|\mathcal{F}\right] \leqslant \left\|\zeta_1\right\|_2^{1+\delta}\left\|\zeta_2\right\|_2^{1+\delta} E\left[\left\|w_i\right\|_2^{1+\delta}|\mathcal{F}\right] E\left[\left\|x_i\right\|_2^{1+\delta}|\mathcal{F}\right] \\ \leq\ & 2^{\delta}\left\|\zeta_1\right\|_2^{1+\delta}\left\|\zeta_2\right\|_2^{1+\delta} E\left[\left\|w_i\right\|_2^{1+\delta}|\mathcal{F}\right]\left(E\left[\left\|v_i\right\|_2^{1+\delta}|\mathcal{F}\right] + \left\|F_T\right\|_2^{1+\delta} E\left[\left\|\Gamma_i\right\|_2^{1+\delta}|\mathcal{F}\right]\right). \end{aligned}$$

The right-hand side is bounded uniformly due to Assumptions 1.ii, iii and v. Therefore, the result follows from the use of the Cramer-Wold device and Theorem A.1 by noting that

$$\begin{aligned} & \frac{1}{N}\sum_{i=1}^{N} w_i' x_i = \frac{1}{N}\sum_{i=1}^{N} w_i' F_T \Gamma_i + \frac{1}{N}\sum_{i=1}^{N} w_i' v_i \\ & \to \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[w_i'|\mathcal{F}\right] F_T \Gamma\left(F_T\right) + \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[w_i' v_i|\mathcal{F}\right] = WX\left(F_T\right) \text{ a.s..} \end{aligned}$$

4. Let ζ_1 and ζ_2 be respectively arbitrary $(h_1+h_3)\times 1$ and $(1+p)\times 1$ vectors. $\frac{1}{N}\sum_{i=1}^{N}\zeta_1' w_i' e_i \zeta_2$ is a sum of $\mathcal{F}$-independent random variables and each term satisfies the following inequality a.s.

$$\begin{aligned} & E\left[\left|\zeta_1' w_i' e_i \zeta_2\right|^{1+\delta}|\mathcal{F}\right] \leqslant \left\|\zeta_1\right\|_2^{1+\delta}\left\|\zeta_2\right\|_2^{1+\delta} E\left[\left\|w_i\right\|_2^{1+\delta}|\mathcal{F}\right] E\left[\left\|e_i\right\|_2^{1+\delta}|\mathcal{F}\right] \\ & \leqslant 2^{\delta}\left\|\zeta_1\right\|_2^{1+\delta}\left\|\zeta_2\right\|_2^{1+\delta} E\left[\left\|w_i\right\|_2^{1+\delta}|\mathcal{F}\right]\left(E\left[\left\|\varepsilon_i\right\|_2^{1+\delta}|\mathcal{F}\right] + \left\|F_T\right\|_2^{1+\delta} E\left[\left\|\gamma_i\right\|_2^{1+\delta}|\mathcal{F}\right]\right). \end{aligned}$$

The right hand side is bounded uniformly due to Assumptions 1.i, iii and iv. Therefore, the result follows since

$$\frac{1}{N}\sum_{i=1}^{N} w_i' e_i = \frac{1}{N}\sum_{i=1}^{N} w_i' F_T \gamma_i + \frac{1}{N}\sum_{i=1}^{N} w_i' \varepsilon_i \to \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[w_i'|\mathcal{F}\right] F_T \gamma\left(F_T\right) \qquad a.s..$$

5. Let ζ_1 and ζ_2 be respectively arbitrary $(h_2+h_4)\times 1$ and $(1+p)\times 1$ vectors. $\frac{1}{N}\sum_{i=1}^{N}\zeta_1{}'x_i{}'e_i\zeta_2$ is a sum of $\mathcal{F}$-independent random variables and each term satisfies the following inequality a.s.

$$\begin{aligned}
&E\left[\left|\zeta_1{}'x_i{}'e_i\zeta_2\right|^{1+\delta}|\mathcal{F}\right]\leqslant\|\zeta_1\|_2{}^{1+\delta}\|\zeta_2\|_2{}^{1+\delta}E\left[\|x_i\|_2{}^{1+\delta}|\mathcal{F}\right]E\left[\|e_i\|_2{}^{1+\delta}|\mathcal{F}\right]\\
&\leqslant\|\zeta_1\|_2{}^{1+\delta}\|\zeta_2\|_2{}^{1+\delta}E\left[(\|v_i\|_2+\|F_T\|_2\|\Gamma_i\|_2)^{1+\delta}|\mathcal{F}\right]E\left[(\|\varepsilon_i\|_2+\|F_T\|_2\|\gamma_i\|_2)^{1+\delta}|\mathcal{F}\right]\\
&\leqslant 4^{\delta}\|\zeta_1\|_2{}^{1+\delta}\|\zeta_2\|_2{}^{1+\delta}\left(E\left[\|v_i\|_2{}^{1+\delta}|\mathcal{F}\right]+\|F_T\|_2{}^{1+\delta}E\left[\|\Gamma_i\|_2{}^{1+\delta}|\mathcal{F}\right]\right)\\
&\cdot\left(E\left[\|\varepsilon_i\|_2{}^{1+\delta}|\mathcal{F}\right]+\|F_T\|_2{}^{1+\delta}E\left[\|\gamma_i\|_2{}^{1+\delta}|\mathcal{F}\right]\right).
\end{aligned}$$

The right-hand side is bounded uniformly due to Assumption 1.i, ii, iv and v. Therefore, the result follows from the use of the Cramer-Wold device and Theorem A.1 by noticing that

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N}x_i{}'e_i&=\frac{1}{N}\sum_{i=1}^{N}v_i{}'\varepsilon_i+\frac{1}{N}\sum_{i=1}^{N}\Gamma_i{}'F_T{}'F_T\gamma_i+\frac{1}{N}\sum_{i=1}^{N}v_i{}'F_T\gamma_i+\frac{1}{N}\sum_{i=1}^{N}\Gamma_i{}'F_T{}'\varepsilon_i\\
&\to\Gamma(F_T)'F_T{}'F_T\gamma(F_T)+\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E[v_i{}'|\mathcal{F}]F_T\gamma(F_T)\ \text{a.s..}
\end{aligned}$$

6. Let ζ_1 and ζ_2 be arbitrary $(1+p)\times 1$ vectors. $\frac{1}{N}\sum_{i=1}^{N}\zeta_1{}'e_i{}'e_i\zeta_2$ is a sum of $\mathcal{F}$-independent random variables and each term satisfies the following inequality a.s.

$$\begin{aligned}
&E\left[\left|\zeta_1{}'e_i{}'e_i\zeta_2\right|^{1+\delta}|\mathcal{F}\right]\leqslant\|\zeta_1\|_2{}^{1+\delta}\|\zeta_2\|_2{}^{1+\delta}E\left[\|e_i\|_2{}^{2+2\delta}|\mathcal{F}\right]\\
&\leqslant 2^{1+2\delta}\|\zeta_1\|_2{}^{1+\delta}\|\zeta_2\|_2{}^{1+\delta}\left(E\left[\|\varepsilon_i\|_2{}^{2+2\delta}|\mathcal{F}\right]+\|F_T\|_2{}^{2+2\delta}E\left[\|\gamma_i\|_2{}^{2+2\delta}|\mathcal{F}\right]\right)\ \text{a.s..}
\end{aligned}$$

The right-hand side is bounded uniformly due to Assumption 1.i and iv. Therefore, we have

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N}e_i{}'e_i&=\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i{}'\varepsilon_i+\frac{1}{N}\sum_{i=1}^{N}\gamma_i{}'F_T{}'F_T\gamma_i+\frac{1}{N}\sum_{i=1}^{N}\varepsilon_i{}'F_T\gamma_i+\frac{1}{N}\sum_{i=1}^{N}\gamma_i{}'F_T{}'\varepsilon_i\\
&\to\Sigma_\varepsilon(F_T)+\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E\left[\gamma_i{}'F_T{}'F_T\gamma_i|\mathcal{F}\right]\ \text{a.s.}
\end{aligned}$$

from the use of Cramer-Wold device and Theorem A.1.

7. Let ζ_1 and ζ_2 be respectively arbitrary $T\times 1$ and $(h_1+h_3)\times 1$ vectors. $\frac{1}{N}\sum_{i=1}^{N}\zeta_1{}'w_i\zeta_2$ is a sum of $\mathcal{F}$-independent random variables and each term satisfies the following inequality a.s.

$$E\left[\left|\zeta_1{}'w_i\zeta_2\right|^{1+\delta}|\mathcal{F}\right]\leqslant\|\zeta_1\|_2{}^{1+\delta}\|\zeta_2\|_2{}^{1+\delta}E\left[\|w_i\|_2{}^{1+\delta}|\mathcal{F}\right].$$

The term on the right-hand side is a.s. uniformly bounded because of Assumption 1.iii. Thus,

$$\text{(B.1.)}\qquad\frac{1}{N}\sum_{i=1}^{N}w_i\to\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E[w_i|\mathcal{F}]\qquad a.s..$$

Notice that

$$\frac{1}{N}\sum_{i=1}^{N}(w_i-\bar{w})'(w_i-\bar{w})=\frac{1}{N}\sum_{i=1}^{N}w_i{}'w_i-\bar{w}'\bar{w}=\frac{1}{N}\sum_{i=1}^{N}w_i{}'w_i-\frac{1}{N}\sum_{i=1}^{N}w_i{}'\frac{1}{N}\sum_{i=1}^{N}w_i.$$

We can use result 1 and (B.1) to conclude that this converges a.s. to

$$W^*(F_T)=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E[w_i{}'w_i|\mathcal{F}]-\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E[w_i{}'|\mathcal{F}]\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}E[w_i|\mathcal{F}].$$

8. Let ζ_1 and ζ_2 be respectively arbitrary $T\times 1$ and $(h_2+h_4)\times 1$ vectors. $\frac{1}{N}\sum_{i=1}^{N}\zeta_1{}'x_i\zeta_2$ is a sum of $\mathcal{F}$-independent random variables and each term satisfies the following inequality a.s.

$$\begin{aligned}
&E\left[\left|\zeta_1{}'x_i\zeta_2\right|^{1+\delta}|\mathcal{F}\right]\leqslant\|\zeta_1\|_2{}^{1+\delta}\|\zeta_2\|_2{}^{1+\delta}E\left[\|x_i\|_2{}^{1+\delta}|\mathcal{F}\right]\\
&\leqslant 2^{\delta}\|\zeta_1\|_2{}^{1+\delta}\|\zeta_2\|_2{}^{1+\delta}\left(E\left[\|v_i\|_2{}^{1+\delta}|\mathcal{F}\right]+\|F_T\|_2{}^{1+\delta}E\left[\|\Gamma_i\|_2{}^{1+\delta}|\mathcal{F}\right]\right).
\end{aligned}$$

The right hand side is bounded uniformly due to Assumption 1.ii and v. Therefore,

$$\text{(B.2.)} \qquad \frac{1}{N}\sum_{i=1}^{N} x_i \rightarrow \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[x_i|\mathcal{F}\right] = F_T\Gamma\left(F_T\right) + \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[v_i|\mathcal{F}\right] \qquad a.s..$$

Write

$$\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\bar{x}\right)'\left(x_i-\bar{x}\right) = \frac{1}{N}\sum_{i=1}^{N} x_i{}'x_i - \bar{x}'\bar{x}.$$

The limit for $\frac{1}{N}\sum_{i=1}^{N} x_i{}'x_i$ is given by result 2. Based on the above, the result follows immediately.

9. Write

$$\frac{1}{N}\sum_{i=1}^{N}\left(w_i-\bar{w}\right)'\left(x_i-\bar{x}\right) = \frac{1}{N}\sum_{i=1}^{N} w_i{}'x_i - \frac{1}{N}\sum_{i=1}^{N} w_i{}'\frac{1}{N}\sum_{i=1}^{N} x_i.$$

The a.s. limits for the three terms are given respectively by result 3, and noting (B.1) and (B.2).

10. Let ζ_1 and ζ_2 be respectively arbitrary $T\times 1$ and $(1+p)\times 1$ vectors. $\frac{1}{N}\sum_{i=1}^{N}\zeta_1{}'e_i\zeta_2$ is a sum of $\mathcal{F}$-independent random variables and each term satisfies the following inequality a.s.

$$\begin{aligned} &E\left[\left|\zeta_1{}'e_i\zeta_2\right|^{1+\delta}|\mathcal{F}\right] \leqslant \left\|\zeta_1\right\|_2{}^{1+\delta}\left\|\zeta_2\right\|_2{}^{1+\delta}E\left[\left\|e_i\right\|_2{}^{1+\delta}|\mathcal{F}\right] \\ &\leqslant 2^{\delta}\left\|\zeta_1\right\|_2{}^{1+\delta}\left\|\zeta_2\right\|_2{}^{1+\delta}\left(E\left[\left\|\varepsilon_i\right\|_2{}^{1+\delta}|\mathcal{F}\right] + \left\|F_T\right\|_2{}^{1+\delta}E\left[\left\|\gamma_i\right\|_2{}^{1+\delta}|\mathcal{F}\right]\right). \end{aligned}$$

The right hand side is bounded uniformly due to Assumption 1.i and iv. Therefore,

$$\text{(B.3.)} \qquad \frac{1}{N}\sum_{i=1}^{N} e_i \rightarrow \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[e_i|\mathcal{F}\right] = F_T\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[\gamma_i|\mathcal{F}\right] = F_T\gamma\left(F_T\right) \qquad a.s..$$

Then the result holds because of result 4, (B.1) and (B.3).

11. Similarly to 10, we can write

$$\begin{aligned} &\frac{1}{N}\sum_{i=1}^{N}\left(x_i-\bar{x}\right)'\left(e_i-\bar{e}\right) = \frac{1}{N}\sum_{i=1}^{N} x_i{}'e_i - \frac{1}{N}\sum_{i=1}^{N} x_i{}'\frac{1}{N}\sum_{i=1}^{N} e_i \\ &\rightarrow \Gamma\left(F_T\right)'F_T{}'F_T\gamma\left(F_T\right) - \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} E\left[\Gamma_i{}'|\mathcal{F}\right]F_T{}'F_T\gamma\left(F_T\right) = 0 \text{ a.s.} \end{aligned}$$

because of result 5, (B.2) and (B.3).

12. Similar to result 10 and 11 of this lemma,

References

Aldous, D. J. and Eagleson, G. K. (1978). On mixing and stability of limit theorems. *The Annals of Probability*, 6(2):325–331.

Chow, Y. S. and Teicher, H. (1997). *Probability theory: independence, interchangeability, martingales.* New York, Springer-Verlag.

Chung, K. L. (1974). *A course in probability theory.* New York ; London, Academic Press.

Daley, D. J. and Vere-Jones, D. (1988). *An introduction to the theory of point processes.* New York, Springer.

Eagleson, G. K. (1975). Martingale convergence to mixtures of infinitely divisible laws. *The Annals of Probability*, 3:557–562.

Forchini, G., Jiang, B., and Peng, B. (2015). Common shocks in panels with endogenous regressors. *Monash Econometrics and Business Statistics Working Papers*, No 8/15.

Hall, P. and Heyde, C. C. (1980). *Martingale limit theory and its application.* New York ; London, Academic Press.

Kuersteiner, G. M. and Prucha, I. R. (2013). Limit theory for panel data models with cross sectional dependence and sequential exogeneity. *Journal of Econometrics*, 174(2):107–126.

Majerek, D., Nowak, W., and Zięba, W. (2005). Conditional strong law of large numbers. *International Journal of Pure and Applied Mathematics*, 20(2):143–156.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of Econometrics. R. F. Engle and D. L. McFadden, Elsevier Science*, 4:2113–2245.

Phillips, P. C. (1988). Conditional and unconditional statistical independence. *Journal of Econometrics*, 38:341–348.

Rao, B. L. S. P. (2009). Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics*, 61:441–460.

Rényi, A. (1963). On stable sequences of events. *Sankhya: The Indian Journal of Statistics, Series A*, 25(3):293–302.

Roussas, G. G. (2008). On conditional independence, mixing, and association. *Stochastic Analysis and Applications*, 26(6):1274–1309.

Stock, J. H. and Wright, J. H. (2000). Gmm with weak identification. *Econometrica*, 68(5):1055–1096.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes with Applications to Statistics*. Springer.

White, H. (2001). *Asymptotic theory for econometricians*. San Diego, Academic Press.